Single Technology Appraisal

Daridorexant for treating insomnia [ID3774]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Daridorexant for treating insomnia [ID3774]

Contents:

The following documents are made available to stakeholders:

Access the **final scope** and **final stakeholder list** on the NICE website.

Pre-technical engagement documents

1. **Company submission from Idorsia Pharmaceuticals Ltd:**

2. **Clarification questions and company responses**

3. **Patient group, professional group, and NHS organisation submissions** from:
   a. The Sleep Charity

4. **External Assessment Report** prepared by Kleijnen Systematic Reviews

5. **External Assessment Report – factual accuracy check**

Post-technical engagement documents

6. **Technical engagement response from company**

7. **External Assessment Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Daridorexant for treating insomnia disorder

ID3774

Document B

Company evidence submission

July 2022

File name	Version	Contains confidential information	Date
ID3774_Daridorexant_Document B_v1.0_AIC_CIC_redacted.docx	Final	No	04/07/2022

Contents

List of Tables

List of figures

Table of abbreviations

AE	Adverse event
AESI	Adverse events of special interest
BAP	British Association for Psychopharmacology
BMI	Body mass index
BWSQ	Benzodiazepine Withdrawal Symptom Questionnaire
BZ	Benzodiazepine
CE	Conformité Européene
CHMP	Committee for Medicinal Products for Human Use
CKS	Clinical Knowledge Summary
CNS	Central nervous system
DB	Double-blind
DISS	Daytime Insomnia Symptom Scale
DORA	Dual orexin receptor antagonist
DSA	Deterministic sensitivity analysis
ECG	Electrocardiogram
EODBT	End-of-double-blind-treatment
EOS	End-of-study
EOT	End of treatment
ER	Emergency room
ESRS	European Sleep Research Society
ESS	Epworth Sleepiness Scale
GABA	Gamma-aminobutyric acid
GDP	Gross domestic product
GLM	Generalised linear model
GP	General practitioner
HCRU	Healthcare resource utilization
HRQoL	Health-related quality of life
ICER	Incremental cost effectiveness ratio
IDSIQ	Insomnia Daytime Symptoms and Impacts Questionnaire
IP	Inpatient
ISB	Independent safety board
ISI	Insomnia Severity Index
LPS	Latency to persistent sleep
LSM	Least squares mean
MedDRA	Medical Dictionary for Regulatory Activities
NHB	Net health benefit
NHS	National Health Service
NHWS	National Health and Wellness Survey
NICE	National Institute for Health and Care Excellence
ONS	Office of National Statistics
OTC	Over-the-counter
PGA-S	Patient Global Assessment of Disease Severity
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity

PICO	Population, intervention, comparator and outcome
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analysis
PSG	Polysomnography
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomized controlled trial
REM	Rapid eye movement
SAE	Serious adverse event
SB	Single-blind
SD	Standard deviation
SDQ	Sleep Diary Questionnaire
SDS	Sheehan Disability Scale
SFIS	Sleep Functional Impact Scale
SLR	Systematic literature review
SmPC	Summary of product characteristics
SUREG	Seemingly unrelated regression
SWS	Slow-wave sleep
TEAE	Treatment-emergent adverse event
TST	Total sleep time
US	United States
VAS	Visual analogue score
WASO	Wake after sleep onset
WPAI	Work Productivity and Activity Index
WTE	Whole time-equivalent

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

This submission addresses the cost-effectiveness, clinical efficacy and safety of daridorexant in adult patients with insomnia disorder. It is aligned with the anticipated marketing authorisation and the final National Institute for Health and Care Excellence (NICE) scope, as outlined in Table 1.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with insomnia disorder	Adults with insomnia disorder	NA
Intervention	Daridorexant	Daridorexant	NA
Comparator(s)	Established clinical management (including sleep hygiene advice) without daridorexant	Established clinical management (including sleep hygiene advice) without daridorexant	NA
Outcomes	The outcome measures to be considered include: • Resolution of symptoms • Changes in sleep patterns and architecture • Sleep quality • Daytime alertness • Recurrence of insomnia • Adverse effects of treatment (including residual daytime sedation and memory impairment) • HRQoL.	The outcomes addressed in this submission include: • Improvement of night-time symptoms of insomnia • Changes in sleep architecture and sleep efficiency • Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function • Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score • Rebound insomnia • Adverse effects of treatment (next-day residual treatment effects and memory	Resolution of symptoms is not an appropriate term to describe the outcome in this submission. The outcome studied in the clinical trials of daridorexant is the quantitative and qualitative improvement of symptoms rather than resolution.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
		impairment) • HRQoL	
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	The cost-effectiveness of daridorexant is presented as cost per QALY. Clinical and cost-effectiveness of the reference case is estimated over a 12-month time horizon.	A short-term model estimating clinical and cost-effectiveness over a 12-month time horizon is presented as the reference case for several reasons. Pharmacodynamics and clinical data of daridorexant demonstrate that the effect of treatment on sleep parameters occurs from the first day of treatment and that the effects on the sleep parameters are mostly lost on the first day of treatment discontinuation. In addition to presenting clinical and cost-effectiveness over a 12-month time horizon, lifetime effects and potential QALY gains from better sleep (e.g., cardiac benefits, reduced fall risk, mortality) is discussed qualitatively in the submission. The potential quantitative impact of having a lifetime model, including impact of improved sleep duration on mortality and the impact of discontinuation, is presented as a scenario.
Subgroups to be considered	The availability and cost of biosimilar and generic products should be considered. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the	None	NA

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	marketing authorisation granted by the regulator.		
Special considerations including issues related to equity or equality	No equity considerations are expected	While digital or face-to-face CBTi is recommended as the first-line treatment for insomnia disorder, it may not be suitable for or accessible to all patients. Daridorexant may thus be suitable for this group of patients as an alternative first-line treatment.	While guidelines recommended CBTi as first-line treatment for insomnia disorder, up to [redacted] of patients refuse CBTi, or cannot access it, when recommended by their GPs. Among those who receive either face-to-face or digital CBTi, [redacted] fail to achieve the desired results, leading to an overall CBTi success rate of only [redacted] (1). Therefore, a broad recommendation for daridorexant to treat insomnia disorder in primary care will provide GPs with a safe and effective option for patients who refuse, cannot access, or fail CBTi.

CBTi = Cognitive behavioural therapy for insomnia; GP = General practitioners; HRQoL = Health-related quality of life; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; NA = Not applicable; NICE = National Institute for Health and Care Excellence; NHS = National Health Services; QALY = Quality-adjusted life year.

B.1.2 Description of the technology being appraised

Daridorexant is a selective and potent dual orexin receptor antagonist (DORA) that decreases wakefulness, thereby allowing a physiological sleep to occur in adult patients with insomnia disorder.

This appraisal considers the proposed indication for daridorexant for the treatment of adult patients with insomnia disorder characterised by symptoms present for at least 3 months and considerable impact on daytime functioning. Table 2 details the technology being appraised in this submission.

Refer to Appendix C for the summary of product characteristics (SmPC) for this technology, pending finalisation of the UK marketing authorisation process.

Table 2: Technology being appraised

UK approved name and brand name	**Daridorexant (QUVIVIQ®)**
Mechanism of action	Daridorexant is a selective and potent DORA, acting as an equipotent orthosteric antagonist at both orexin 1 and orexin 2 receptors, with equilibrium dissociation constants (K_b) of these antagonisms of 0.5 nM and 0.8 nM in humans, respectively (2). The orexin neuropeptides (orexin A and orexin B) act on orexin receptors to promote wakefulness. Daridorexant blocks the binding of orexin neuropeptides to the receptors and consequently decreases the wake-drive, allowing sleep to occur (3). As a DORA, daridorexant acts by decreasing wakefulness, which contrasts with the mechanism of action of sedative/hypnotic medications (such as benzodiazepines and Z-drugs) that induce sleep through general suppression of the CNS via GABA-A receptor agonism (3-5).
Marketing authorisation/CE mark status	Currently, daridorexant does not have a UK marketing authorisation. In March 2021, marketing authorisation application for daridorexant was submitted to the EMA. A positive CHMP opinion was issued in February 2022, and it was approved on 2nd May 2022 by EMA (6). A marketing authorisation submission has already been made to MHRA.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Daridorexant is indicated for the treatment of adult patients with insomnia characterised by symptoms present for at least 3 months and considerable impact on daytime functioning (3). The SmPC is provided in Appendix C.

Method of administration and dosage	Daridorexant is available as 25 mg and 50 mg film-coated tablets (3). The recommended dose for adults is one tablet of 50 mg once per night, taken orally in the evening within 30 minutes before going to bed. Based on clinical factors (moderate hepatic failure or concomitant use of moderate CYP3A4 inhibitor medicines), some patients may be treated with 25 mg once per night. The maximum daily dose is 50 mg (3). The treatment duration should be as short as possible. The appropriateness of continued treatment should be assessed within 3 months and periodically thereafter (3). The UK SmPC (which is the current EU SmPC) is provided in Appendix C.
Additional tests or investigations	No additional tests or investigations are required for identification of the population for daridorexant administration.
List price and average cost of a course of treatment	[redacted]/ day
Patient access scheme (if applicable)	Not applicable.

CHMP=Committee for Medicinal Products for Human Use; CNS = Central nervous system; CYP3A4 = Cytochrome P450 3A4; DORA = dual orexin receptor antagonist; EMA=European Medicines Agency; GABA-A = Gamma-aminobutyric acid Type A; MHRA= Medicines and Healthcare products Regulatory Agency; PSG = polysomnography; SmPC=summary of product characteristics; UK=United Kingdom.

B.1.3 Health condition and position of the technology in the treatment pathway

- Patients with insomnia suffer from both night-time symptoms and daytime functioning impairment, affecting subjective and objective dimensions of health (7).
- Insomnia presents as either acute or chronic insomnia. Chronic insomnia, also known as insomnia disorder, is defined as symptoms occurring for ≥3 nights per week for ≥3 months together with daytime impairment (8, 9).
- Patient-reported sleep disturbances and impaired daytime functioning are the cornerstone for diagnosing and managing insomnia disorder in primary care clinical practice. This subjective approach of diagnosis is similar to mental health conditions such as depression and schizophrenia (9, 10).

- Primary care clinicians can utilise insomnia-specific patient-reported outcome (PRO) instruments to rapidly assess symptoms and severity of insomnia disorder and ensure optimal treatment for eligible patients.

B.1.3.1 Disease overview

In general practice, clinicians frequently encounter patients with sleep disorders related to sleep initiation or maintenance, often referred to as insomnia (11). Patients with insomnia suffer from both night-time symptoms and daytime functioning impairment, affecting subjective and objective dimensions of health (12, 13). Insomnia impairs an individual's overall quality of life (QoL) by affecting physical and mental health, cognitive ability, mood, and behaviour. Moreover, impaired daytime functioning due to insomnia increases the risk of errors or accidents, impacts task performance and negatively impacts social life, relationships, and family life (14). If left untreated, insomnia can increase the risk of depression, anxiety, and in the long term, diabetes and cardiovascular disease (CVD) (14).

Definition of insomnia disorder

Insomnia disorder is characterised by type, frequency and duration of symptoms using the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; 2013) and the International Classification of Sleep Disorders, Third Edition (ICSD-3; 2014) diagnostic classification systems:

- Insomnia is defined by the DSM-5 (2013) as dissatisfaction with sleep quantity or quality that is accompanied by one (or more) of the following symptoms: difficulty initiating and maintaining sleep, characterised by frequent awakenings or problems returning to sleep, and early morning awakenings with inability to return to sleep. The sleep disturbance causes clinically significant distress or impairment, with detrimental effects on daytime functioning, including social, occupational, educational, academic, behavioural, or other important areas of functioning (15).

- The ICSD-3 (2014) defines insomnia as persistent difficulty with sleep initiation, duration, consolidation or quality that occurs despite the opportunity and circumstances for sleep, resulting in some daytime functional impairment (8). This

definition of insomnia disorder is consistent with that of the International Classification of Diseases, 11th Revision (8, 16).

Based on the duration and frequency of sleep disturbance, DSM-5 and ICSD-3 further classify insomnia as acute and chronic insomnia (8, 15):

- Acute or short-term insomnia refers to transient episodes of insomnia, with symptoms lasting <3 months, that usually occur after an emotional or psychological stress and resolve quickly once the underlying causes are addressed.
- Chronic insomnia (also referred to as insomnia disorder in this submission) is defined as insomnia symptoms occurring for ≥3 nights per week for ≥3 months.

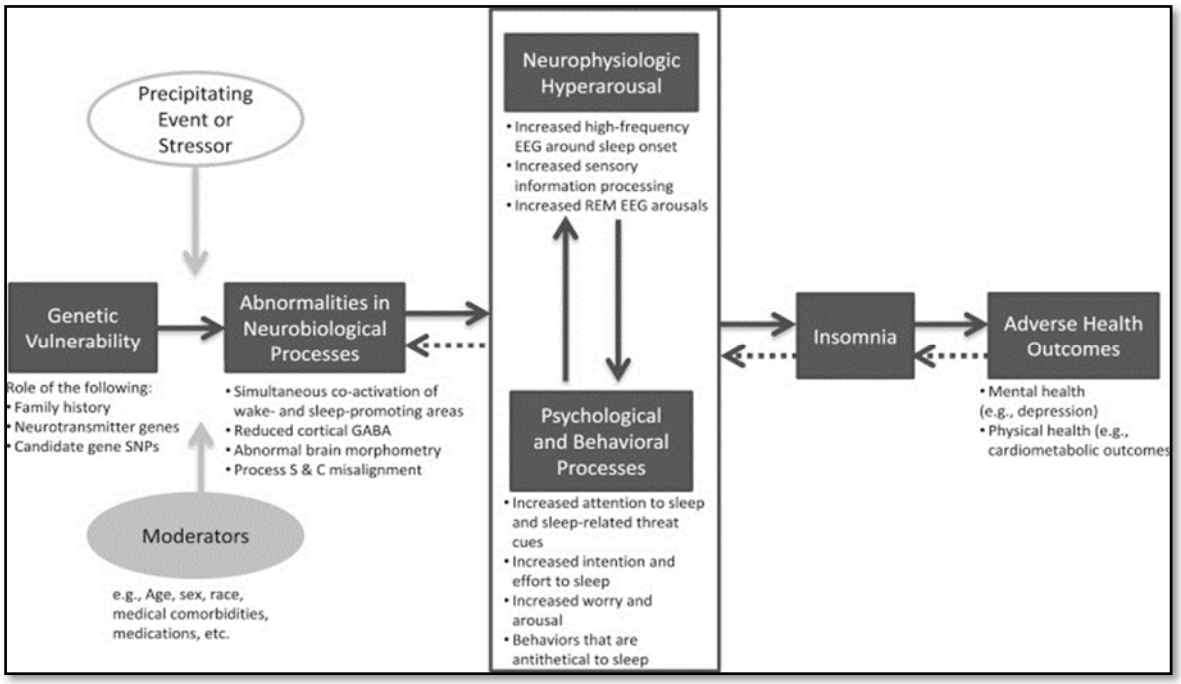
Pathogenesis

The sleep-wake cycle is regulated by homeostatic and circadian processes. Together with a complex interplay of molecular, genetic, neurological and psychological factors, these processes achieve balance between sleep and wake (17). Therefore, any impact on these factors is anticipated to influence the sleep-wake cycle, resulting in insomnia in some cases.

The pathophysiology of insomnia disorder is multifaceted that involves molecular anomalies, genetic polymorphisms, gene-environment interactions, and life-events (18, 19). Molecular anomalies include dysregulation of neurotransmitters involved in the sleep–wake cycle, such as gamma-aminobutyric acid (GABA), melatonin, orexin, norepinephrine, histamine, serotonin, acetylcholine, and dopamine (18). Genetic factors implicated in the predisposition to insomnia disorder have been elaborated upon in genome-wide association studies (20, 21). Polymorphisms in the serotonin-transporter-linked promoter region (5-HTTLPR), protein-coding period family of regulatory genes of the circadian rhythm (PER2/3), apolipoprotein E4 (APOE4) and peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) genes of the dopaminergic system increase the risk of insomnia. The risk is further amplified by gene-environment interactions such as job stress or alcoholism (22). Life events such as trauma, lifestyle factors, psychological processes and behaviours can each play precipitating and perpetuating roles in the aetiology of insomnia disorder.

Figure 1 illustrates an integrated model of the development of insomnia disorder through the interplay between predisposing and moderating factors (e.g., genetic vulnerability, character traits or medical conditions) and the presence of precipitating occurrences (e.g., major life events, stress or trauma) (23).

Figure 1: Theoretical model of the pathophysiology of insomnia disorder (23)



EEG = electroencephalogram; GABA = gamma-aminobutyric acid; REM = rapid eye movement; SNP = single nucleotide polymorphism.
Note: Permission must be sought from the publisher before reproducing this figure for use with an external audience.

Clinical presentation

Patients affected by insomnia disorder present with symptoms of sleep disturbances and impaired daytime functioning (24). Clinicians often record these symptoms based on patients' medical/clinical history and patient-reported measures. Primary drivers of patients' perception of their insomnia include (25):

- Sleep symptoms such as difficulty falling asleep, difficulty maintaining sleep and undesired early-morning awakening.
- Symptoms of daytime impairment, including somnolence, fatigue, malaise, irritability, concentration and memory impairment.

Overall, experience of sleep is an important factor to consider when assessing insomnia symptoms and, while guidelines define daily sleep needs, optimal sleep

varies between individuals based on their circadian rhythm (24, 26). This subjective assessment and management approach is consistent with other psychological conditions, such as depression, where diagnosis and management are based on subjective reports by patients and/ or their caregivers (27). This contrasts with chronic conditions, such as hypertension or diabetes mellitus, which rely on objective measures (9, 10). Despite being subjective, patient-reported symptoms are the key to diagnosing and managing insomnia disorder, justifying that subjective assessment is clinically as important as objective measures of health.

Diagnosis

The DSM-5 and the ICSD-3 recommends that insomnia should be considered as an independent disorder, regardless of the presence of comorbidities or any ongoing physical or mental disorder, including any concurrent sleep disorder (8, 15). An independent diagnosis of insomnia disorder should be made, taking into consideration the type, duration and frequency of symptoms. It is also recommended to consider the temporal relationship between the development of insomnia symptoms and those of other medical conditions, since insomnia disorder can precipitate or worsen existing comorbidities (8, 15).

Most patients present with symptoms of “trouble sleeping” in primary care, either at first onset or when it has a substantial impact on their daytime functioning. General practitioners (GPs) diagnose insomnia disorder based on patient self-reports of insomnia symptoms (see *Clinical presentation*) and correlate them with the diagnostic criteria of the DSM-5 or ICSD-3 (8, 15).

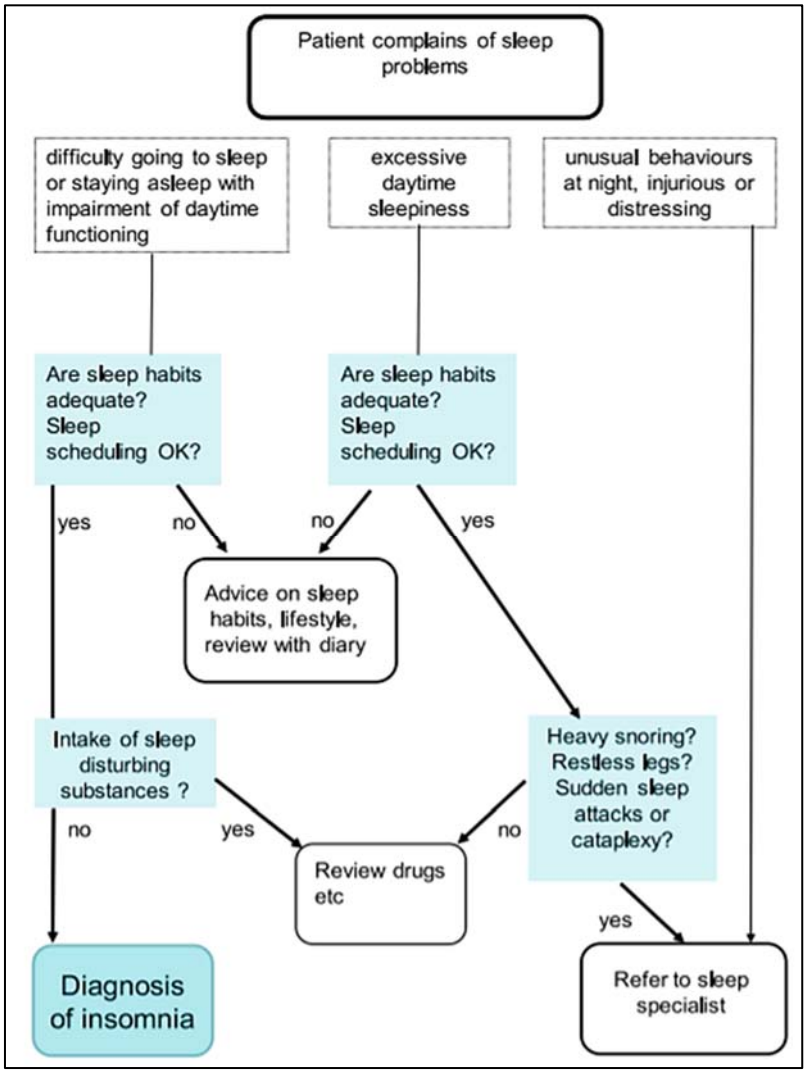
The British Association for Psychopharmacology (BAP) guidelines recommend comprehensive assessment of subjective insomnia symptoms (Figure 2) (14):

- Assessment of patient’s current sleep habits, including symptoms of difficulty getting sleep and/or staying asleep, frequency of this occurrence, its persistence and daytime effect, with the use of a clinical rating scale, such as a 2-item Sleep Condition Indicator (SCI) (28), or a sleep diary to assess sleep difficulties over time and gauge the potential contribution of poor sleep and lifestyle habits to daytime impairment,

- Assessment of sleep history, including former/present sleep disorders, information from a bed partner, circadian factors and sleep-wake factors,

- Medical history, medication and substance use, physical examination and additional measurements as needed, assessment of psychiatric and psychological history, personality factors, and occurrences or conflicts in work and personal life.

Figure 2: Diagnosis of insomnia (14)



It is important to note that although polysomnography (PSG) is the gold standard of objective sleep assessment it is not recommended for diagnosis of insomnia or its routine evaluation. PSG and actigraphy, another objective measure, is only indicated if other sleep disorders such as sleep apnoea or narcolepsy are suspected. In case of other sleep disorders, a referral to specialty care should be made for further investigation and management of symptoms (14).

Assessment of insomnia disorder

As per the DSM-5 definition of insomnia disorder, it is essentially a subjective patient experience wherein PRO instruments may play an important role in its diagnosis and management. Currently available insomnia-specific PRO instruments assess patients' sleep habits and their impact on daytime functioning impairment. These include the Daytime Insomnia Symptom Scale (DISS), (29) the Daytime Consequences of Sleep Questionnaire (DCSQ), (30) the Functional Outcomes of Sleep Questionnaire (FOSQ), (31) the Pittsburgh Insomnia Rating Scale (PIRS), (32) the Profile of Mood States (POMS), (33) the Sleep Functional Impact Scale (SFIS), the Insomnia Severity Index (ISI) (34), the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) (35) and the Epworth Sleepiness Scale (ESS) (36). Other PRO instruments that focus on night-time symptoms include the SCI (28).

Numerous PRO instruments are available to assess patients' perception of their sleep problems and the associated impact on daytime functioning. Among them, the PROs described below were part of daridorexant clinical development program.

Assessment of insomnia symptoms:

1) The Sleep Diary Questionnaire (SDQ), derived from the Consensus Sleep Diary assessed sleep quantitatively (subjective TST, subjective WASO, subjective LSO), and qualitatively (sleep depth and sleep quality) along with morning sleepiness and the ability to function during the day.

2) The IDSIQ, developed and validated to assess daytime functioning impairment in insomnia.

Global assessment of insomnia severity:

3) Insomnia Severity Index© (ISI©).

4) Sheehan Disability scale© (SDS©).

Clinical guidelines do not recommend the use of any specific PRO instrument to assess insomnia symptoms in clinical practice (14). Although some PRO instruments are validated for use in primary care, no single instrument is recommended or more

widely used than the other (37). Primary care clinicians can make use of an insomnia PRO instrument, which has been validated in clinical practice to rapidly assess the symptoms and severity of insomnia disorder to ensure optimal treatment for eligible patients. Ideally, a validated PRO instrument should be suitable for primary care use, should be easy to understand, quick to administer (e.g., <3 minutes), assess both daytime and night-time symptoms, be available in English, and administered by non-physicians.

Generic preference-based PRO instruments, such as the EuroQol-5D (EQ-5D), have been widely used across different patient groups and indications to ensure comparability between interventions (38). Yet, a key concern with these instruments is that they may be missing relevant or important dimensions for some specific conditions. Specifically, the EQ-5D does not include a fatigue dimension that is highly relevant to insomnia (39). This is discussed in greater detail in B.3.13 Benefits not captured in the QALY calculation. Therefore, generic preference-based PRO instruments may not be sensitive to changes when used to monitor response to insomnia treatment or disease progression.

A novel mapping algorithm was applied to derive EQ-5D utilities from the ISI© scores that was used as a key effectiveness parameter for the cost-effectiveness model (40). In this submission, the derived EQ-5D utilities were incorporated into the model presented in B.3 Cost effectiveness.

Insomnia Severity Index©

The ISI© assesses subjective symptoms, as well as the degree of concern or distress caused by the symptoms and their consequences. ISI© comprises of seven items which measures patients' perception of insomnia severity in the past two weeks (34).

Individual items assess the severity of sleep-onset and sleep maintenance difficulties, as well as satisfaction with current sleep patterns, interference with daily functioning, noticeability of impairment attributed to the sleep problem, and the degree of distress or concern caused by the sleep problem. Each item of the ISI© is rated on a 0 to 4 scale, and the total score ranges from 0 to 28. A higher score suggests more severe insomnia (34). The total score is interpreted as follows: absence of insomnia (0–7), subthreshold insomnia (8–14), moderate insomnia (15–21), and severe insomnia (22–

28) (41). Thus, the instrument can be used to identify patients with moderate-to-severe insomnia disorder. In the phase III daridorexant clinical trials, ISI© was used to screen patients with moderate-to-severe insomnia disorder for enrolment in the confirmatory study.

ISI© can be easily administered, with scores that can be calculated in less than a minute. Its validity has been demonstrated in the primary care setting (37). Therefore, it can be potentially used to identify patients with moderate-to-severe insomnia disorder and ensure they receive optimal treatment. In addition, ISI© can also be used to monitor patients' progress with therapy in clinical practice.

Insomnia Daytime Symptoms and Impacts Questionnaire

The IDSIQ is a PRO instrument developed by modifying the DISS following interviews with patients with insomnia and discussions with experts in insomnia research (35). The instrument is administered daily in the evening and has a recall period of "today" and comprises of 14 items structured across 3 domains of alert/cognition (6 items), mood (4 items) and sleepiness (4 items). These domains reflect the commonly encountered daytime functioning effects of insomnia. Each question is scored from 0 to 10, and the domain scores are summed, with higher scores indicating worse symptoms and impact of insomnia (35).

The IDSIQ was developed for use in clinical studies as the first PRO instrument to assess impairment of daytime functioning due to insomnia. It was validated according to the United States Food and Drug Administration guidance for industry (42), demonstrating strong internal consistency (Cronbach's alpha: 0.917 for IDSIQ total score, 0.806–0.918 for domain scores) and test-retest reliability (intra-class correlation coefficient: 0.856–0.911) (35). Meaningful change thresholds include 17 for IDSIQ total score, 9 for the alert/cognition domain, 4 for the mood domain and 4 for the sleep domain (43). Currently, IDSIQ has been used as a key secondary endpoint in the phase III confirmatory study and as an exploratory endpoint in the extension trials of daridorexant. Its validity and feasibility for use in the primary care setting have not been established.

Sheehan Disability Scale©

The Sheehan Disability Scale (SDS©) is a validated instrument with 5 items; 3 scales presented visually as a horizontal line marked with both numbers (0 to 10) and verbal anchors (0 = Not at all, 1–3 = mildly, 4–6 = moderately, 7–9 = markedly, and 10 = extremely) to score the disruption of the symptoms on 1= work/school work; 2 = social life/leisure activities; 3 = family life / home responsibilities; two questions with the last week as a recall period on: 4 = days missed in the past week; 5 = days underproductive in the past week. Two outcomes can be calculated 1) absenteeism = Q4; 2) presenteeism=Q1/10*Q5 (44, 45). The SDS© was a safety outcome in the clinical trials of daridorexant as described in B.2 Clinical effectiveness.

B.1.3.2 Epidemiology

- The prevalence of insomnia disorder in the general population is increasing due to an ageing population and an increasing burden of mental health conditions and chronic diseases (46).
- In England, 9.3 million adults, or one in every five, are estimated to experience insomnia symptoms, which is comparable to other mental health conditions such as depression (47).
- Approximately 3.3 million adults, or one in every three (35%) with insomnia symptoms, meet the DSM-5 criteria for insomnia disorder (47).
- Multiple psychiatric and medical conditions are frequently associated with insomnia and may have a reciprocal relationship (18, 48).
- Insomnia disorder is a common comorbid condition and a risk factor for chronic diseases and mortality (12, 49, 50).
- Despite the associated comorbidities, primary care guidelines direct towards independent diagnosis and treatment of insomnia disorder (14, 51).

Incidence

The incidence of insomnia disorder in the UK is poorly characterized. A population-based, longitudinal cohort study using postal self-answered questionnaires at baseline and 12-month follow-up was conducted in the UK among adults aged ≥18 years (N=4,885) (52). Among 859 respondents who did not report insomnia at baseline and responded to the sleep questions at follow-up, 125 were diagnosed with insomnia at 1 year. This corresponds to an annual incidence of 14.6% (95% confidence interval 12.2–16.9), or 7.5 million adults when extrapolated to the English population (52).

Prevalence

Only a few studies have estimated the prevalence of insomnia disorder using the current DSM-5 criteria (14, 53-56). No recent UK studies estimating the prevalence of DSM-5 insomnia disorder in the general population were identified. Therefore, data from the 2020 National Health and Wellness Survey (NHWS) were utilised to calculate the prevalence of insomnia disorder in the UK (57, 58).

The Cerner Enviza NHWS is a large, nationally representative, self-administered, internet-based questionnaire involving adults (aged 18 years or older) in the US, UK, France, Germany, Italy, Spain and Japan (57, 58). It is projected to reflect the general population of the country surveyed using known population incidences for key subgroups. Potential respondents to the NHWS are recruited through an existing, general-purpose Web-based consumer panel. The consumer panel recruits panel members through opt-in e-mails, co-registration with panel partners, e-newsletter campaigns, banner placements, and affiliate networks. The NHWS contains data on over 1 million patients globally to date (57, 58). As the NHWS is a nationally representative and widely used dataset in publications and other company submissions, it was used as the primary source of insomnia disorder prevalence in the budget impact analysis of this submission.

The 2020 NHWS data included 10,408 respondents in the UK with a mean age of 47.3 years and 53.3% were females (n=5,544). Approximately 1 in 5 respondents (21%) reported experiencing symptoms of insomnia. When extrapolated to the adult population in England (44.4 million), this corresponds to 9.3 million people experiencing insomnia symptoms (59). The prevalence of insomnia disorder,

estimated based on the number of respondents who experienced insomnia symptoms at least four times per week and for at least 6 months in the past year, was 7.3% (47).

Individuals with insomnia disorder tend to feel anxious and helpless over their inability to sleep or lack of sleep and their increasing feelings of isolation, leading to the development of depression and anxiety disorders. Conversely, the presence of depression and anxiety can cause sleepless nights resulting in insomnia disorder, suggesting an overlap of symptoms and correlation between insomnia disorder and psychological conditions (60). As the population continues to age, the prevalence of insomnia disorder and the closely related psychological conditions are both projected to increase (46). In the UK, 1 in 5 adults aged ≥16 years, experience depressive symptoms, which is similar to the current prevalence of insomnia (47, 61). Therefore, prioritising the improvement of outcomes in patients with insomnia disorder is just as important as for other burdensome psychological conditions.

Risk factors

The multifactorial aetiology of insomnia involves a range of risk and lifestyle factors, including demographics, clinical, and genetic factors (18, 48). Multiple psychiatric and medical conditions are also frequently associated with or may have reciprocal relationship with insomnia disorder. Approximately 50% of patients with insomnia also have mood (e.g., major depressive disorder) or anxiety disorders (e.g., PTSD) (19, 25). In the UK, 68% of insomnia patients experience depression, while 75% experience anxiety (62).

Insomnia disorder is a common comorbid condition and a risk factor for chronic diseases and mortality. A meta-analysis of 14 prospective cohort studies demonstrated an increased risk of hypertension in insomnia patients compared to those without insomnia (RR=1.21, [95% CI 1.10, 1.33]) (49). A cross-sectional study involving 1,311 insomnia sufferers in Belgium reported that those with <6.5 hours of sleep had higher odds of having type 2 diabetes compared to those with ≥8 hours of sleep (OR=1.81, [95% CI 1.15, 2.84]) (50).

A systematic review and meta-analysis of prospective cohort studies assessed the relationship between sleep duration, all-cause mortality and cardiovascular events (12). A pooled analysis of 57 studies showed that when sleep duration was <6 hours

a day, all-cause mortality increases by 6% (95% CI [4 to 7%]) per 1-hour reduction in sleep duration (12). Similarly, an analysis of 37 studies reported a 6% (95% CI [3 to 8%]) increase in risk of cardiovascular events per 1-hour reduction in sleep duration (12). The findings from these studies demonstrate the increased risk of chronic diseases and mortality associated with insomnia disorder.

Despite the associated comorbidities, BAP and NICE guidelines advocate for independent diagnosis and treatment of insomnia disorder in primary care (14, 51). Referral to a specialist should only be considered if other sleep disorders are suspected or if treatment in primary care has failed.

B.1.3.3 Humanistic burden

- Insomnia disorder has a substantial impact on patients' QoL, affecting their physical, social, emotional and psychological well-being (63).
- The impact of insomnia disorder on patients' QoL is significant regardless of diagnosis or treatment status (47).

Overall, insomnia disorder has a significant impact on patients' physical, mental, and social well-being. Insomnia disorder is associated with considerable daytime impairment, often leading to daytime sleepiness, fatigue, and irritability which contributes to depression and anxiety. It can impact patients' ability to complete tasks and participate in activities of daily living, and increase the risk of home, work and motor vehicle accidents. Insomnia disorder is also associated with a higher risk of falls and injuries among the elderly (63). Globally, at least 10% of insomnia patients experience daytime impairment due to its symptoms (64-66). In the UK, 52% of patients reported severe impact on personal and professional life due to insomnia (62).

The systematic literature review (SLR) of the humanistic burden of insomnia disorder (Appendix H) did not identify any studies conducted in the UK. Therefore, the negative impact of insomnia disorder on HRQoL was quantified using the UK subset of the 2020 NHWS data (N=10,408) (47). Mean [SD] EQ-5D utility of 2,128 respondents who reported insomnia symptoms was significantly lower than that of the general cohort (n=8,280) (0.68 [0.25] vs 0.82 [0.21], p<0.001). Subgroup analysis showed that the

impaired HRQoL was apparent regardless of insomnia disorder diagnosis or treatment status.

The growing evidence on the health and QoL impacts of insomnia has prompted the UK government to review its policy on sleep in care settings. In its 2020 prevention green paper, the government planned to review the evidence on sleep and health and determine steps to ensure those in care settings are getting sufficient rest (67).

B.1.3.4 Economic burden

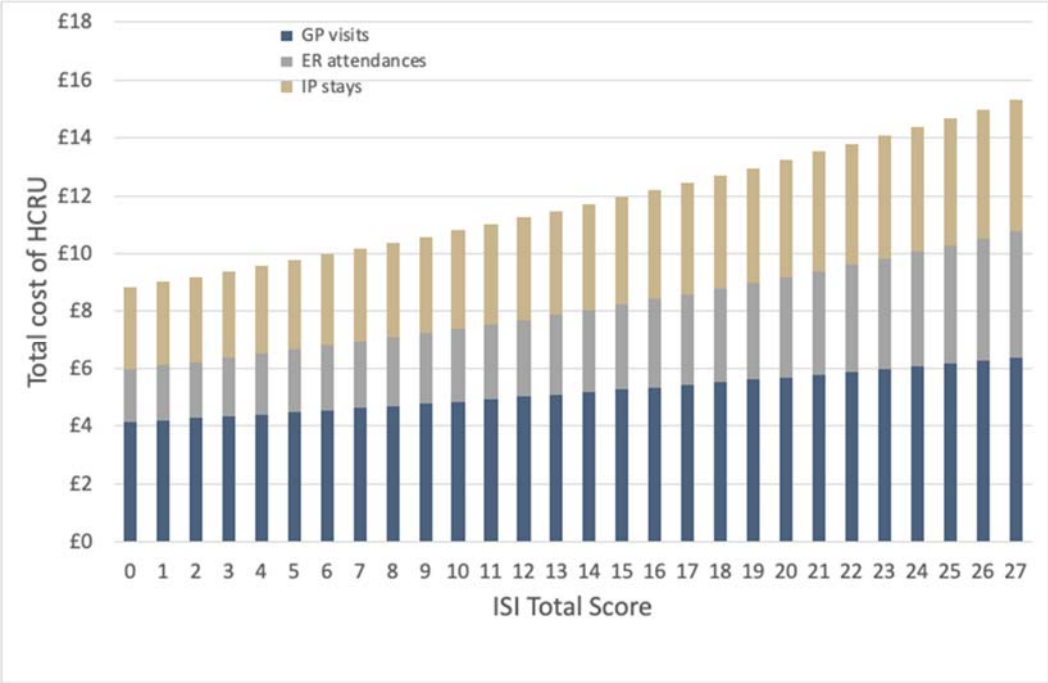
- Economic loss associated with insufficient sleep is estimated to account for 1.36%–1.86% of the UK's gross domestic product (GDP) (68).
- Direct medical costs increase with worsening insomnia disorder (47).
- Productivity loss is the primary driver of the economic burden of insomnia disorder, amounting to £41.2 billion lost each year from lost and unproductive workdays (69, 70).
- The use of hypnotics beyond the recommended duration for insomnia disorder (<4 weeks) may exacerbate productivity loss due to somnolence and impaired daytime functioning (71, 72).

The economic burden of insomnia disorder is challenging to quantify due to the subjective nature of the condition and its frequent occurrence together with other conditions that confer a much higher burden (19, 25). The RAND 2016 economic predictions estimated the economic loss of insufficient sleep at 1.36%–1.86% of the UK's GDP. This is forecasted to increase to 1.63%–2.17% by 2030. Thus, reduced productivity and direct medical costs related to insufficient sleep are associated with substantial economic losses in the UK (68).

The SLR of the economic burden of insomnia disorder (Appendix I) did not identify any studies conducted in the UK. Therefore, economic burden of insomnia disorder was quantified using data from the NHWS. Analysis of the UK subset of NHWS 2020 data (N=10,408) showed that healthcare resource utilisation (HCRU) increased with

increasing ISI© score (Figure 3), indicating that economic burden increases with worsening insomnia disorder (47).

Figure 3: Total predicted direct healthcare costs by ISI© score (shown as the value of a one-point reduction to that score) (47)

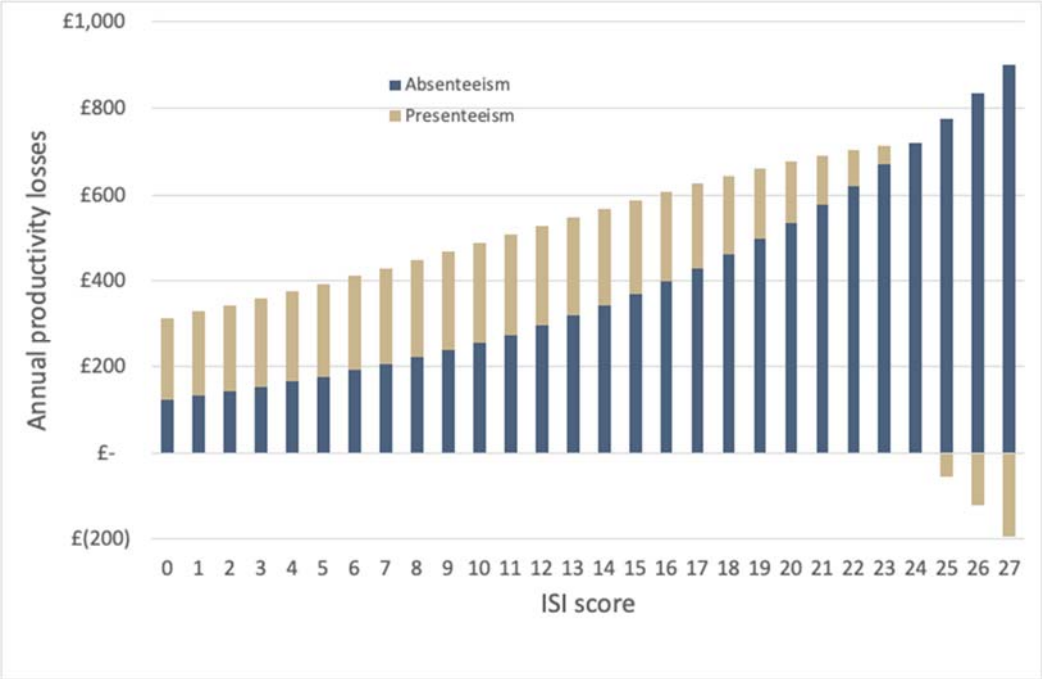


ER = Emergency room; GP = General practitioners; IP = Inpatient; ISI = Insomnia Severity Index.

Productivity loss is the main driver of the economic burden of insomnia disorder (69, 70). Data from the Work Productivity and Activity Index (WPAI) questionnaire administered in NHWS 2020 was used to estimate productivity loss due to insomnia. A higher WPAI score indicates greater work and activity impairment. Respondents in the UK subset (N=10,408) who reported insomnia symptoms had greater work productivity (31.7 [29.9] vs 20.6 [28.9], p<0.001) and activity impairment (35.5 [29.1] vs 20.3 [26.7], p<0.001) compared to those who did not (47). Further analysis of the WPAI and ISI© data demonstrated increasing absenteeism and presenteeism with an increasing ISI© score (Figure 4), indicating that more severe insomnia symptoms is associated with greater productivity loss.

The productivity loss estimated using NHWS 2020 translates to an annual loss of 52 days of work and up to 4 months spent being unproductive at work due to insomnia symptoms. This amounts to an annual productivity loss of £41.2 billion when extrapolated to the 3.3 million adults suffering from insomnia disorder in England (47), highlighting the substantial economic burden associated with the condition.

Figure 4: Estimated productivity losses due to absenteeism and presenteeism as a function of ISI© (shown as the value of a one-point reduction to that score) (47)



ISI = Insomnia Severity Index.

In the UK, approximately 300,000 patients are on short-term therapies for 12 months or longer (62). Chronic use of hypnotics is associated with somnolence and impaired daytime functioning, further adding to the already substantial impact of insomnia disorder on work productivity (71, 72). Due to concerns with chronic use of hypnotics resulting in tolerance or dependence, a recent NICE guidance (NG215) was published to provide recommendations on how and when these therapies should be recommended to patients (73). Specifically, NG215 states that the duration of therapy should reflect the management plan and comply with guideline recommendations, and recommends prescribers to communicate clearly to patients about the intended duration of therapy and plans for periodic review.

In addition, the Sheehan Disability Scale© (SDS©) utilized in the phase III trials of daridorexant was used to estimate productivity loss (44, 45). The SDS© is described as a safety outcome in B.2 Clinical effectiveness and estimation of productivity loss is presented as a scenario analysis in B.3.5.4 Miscellaneous unit costs and resource use.

B.1.3.5 Clinical care pathway

- The majority of patients with insomnia disorder are treated in primary care (47). Specialist referral is made only in the case of a suspected sleep disorder or mental health issue.
- The goals of therapy in insomnia disorder are to improve sleep, reduce suffering and improve daytime functioning (14, 51).
- None of the existing guidelines include recommendations for long-term treatment of insomnia disorder, mainly due to the safety limitations of the current treatment options.
- Current guidelines recommend cognitive behavioural therapy for insomnia (CBTi) as first-line therapy (14, 25, 51). Sleepio®, a self-help digital sleep improvement programme based on CBTi, is recommended by NICE as a cost-saving option for treating insomnia and insomnia symptoms (MTG70) (74).
- BZs and Z-drugs are recommended as efficacious hypnotics for insomnia disorder, but should be restricted to the shortest possible duration (<4 weeks) due to the safety risks associated with prolonged usage (14, 25, 51). Use of prolonged-release melatonin in patients aged ≥55 years is recommended for an initial duration of 3 weeks, as it improves sleep-onset latency and quality. If an adequate response is achieved, the treatment can be continued for another 10 weeks (75).
- Existing pharmacotherapies are prescribed for longer-than-recommended durations due to the lack of safe and effective treatment alternatives for insomnia disorder (76, 77).

Patient journey

The typical journey of insomnia patients is summarised below (an illustration is presented in Figure 5) (62). With onset of insomnia symptoms, patients self-initiate healthy behavioural changes, including sleep hygiene practices, diet and exercise.

Additionally, patients take plant-based products or over-the-counter (OTC) antihistamines based on their own research or under the advice of friends/family or pharmacist. The benefits of these approaches are often transient.

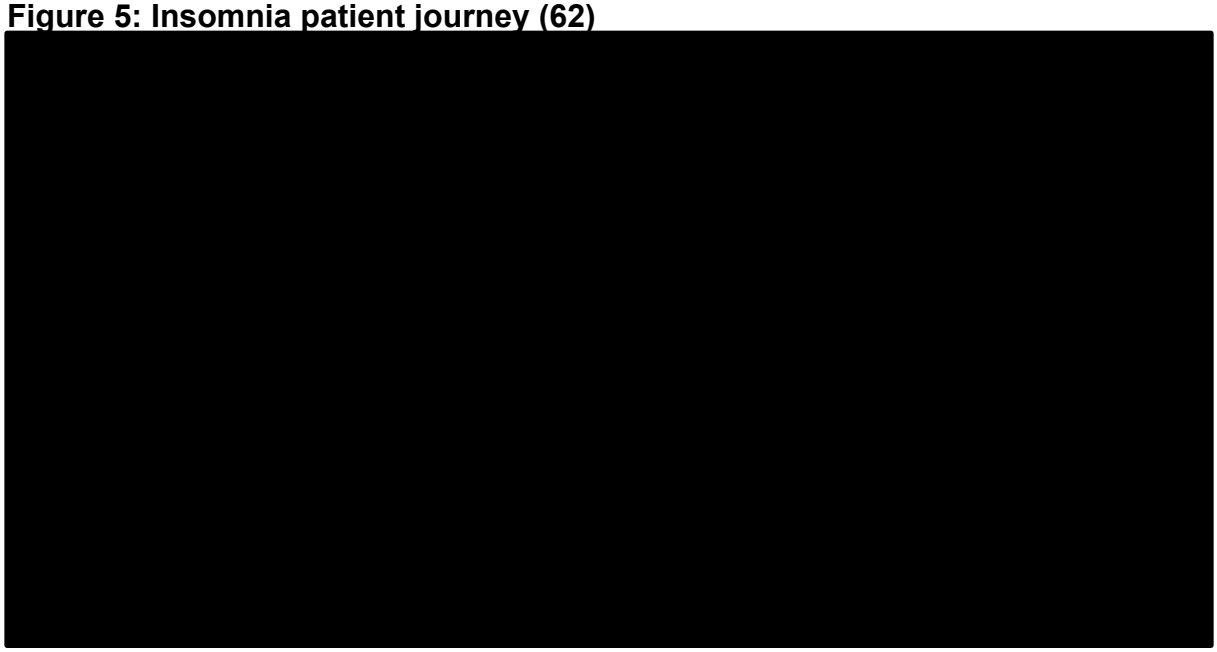
When symptoms are severe enough to affect patients' daily activities, they consult their GPs. A survey conducted in May 2022 among more than 1,000 GPs in the UK reported that [redacted] (1). [redacted] (1). Pharmacotherapies are recommended only for a short duration (<4 weeks or <13 weeks for melatonin) with the aim of re-establishing natural, subjectively satisfying sleep (14, 51).

Hypnotics can effectively treat night-time symptoms of insomnia (such as sleep onset and/or sleep maintenance). However, psychological dependence often leads to the use of these medications beyond the recommended duration, as most patients are willing to trade-off the safety concerns for a good night's sleep (72). Considering insomnia disorder is a pervasive condition with a chronic course, most patients continue short-term treatments for extended periods despite safety concerns such as risk of dependence, next-day residual effects and withdrawal symptoms or rebound upon discontinuation. Insomnia being a refractory condition reoccurring upon treatment discontinuation, GPs tend to continue these treatments to manage patient expectations despite being aware that they are only indicated for short-term use. [redacted] (76, 77). Due to the prolonged usage of hypnotics and other medicines associated with dependence or withdrawal symptoms, coupled with the lack of alternative treatment options, NICE recently issued a guidance (NG215) to provide recommendations on safe prescription and withdrawal management of these drugs (73).

The UK GP survey revealed an [redacted], suggesting that majority of the patients are managed in primary care. This may be due

to under-developed referral pathways and referrals being considered only in cases of other suspected disorders or if the patient is in occupational at-risk group (1).

Figure 5: Insomnia patient journey (62)



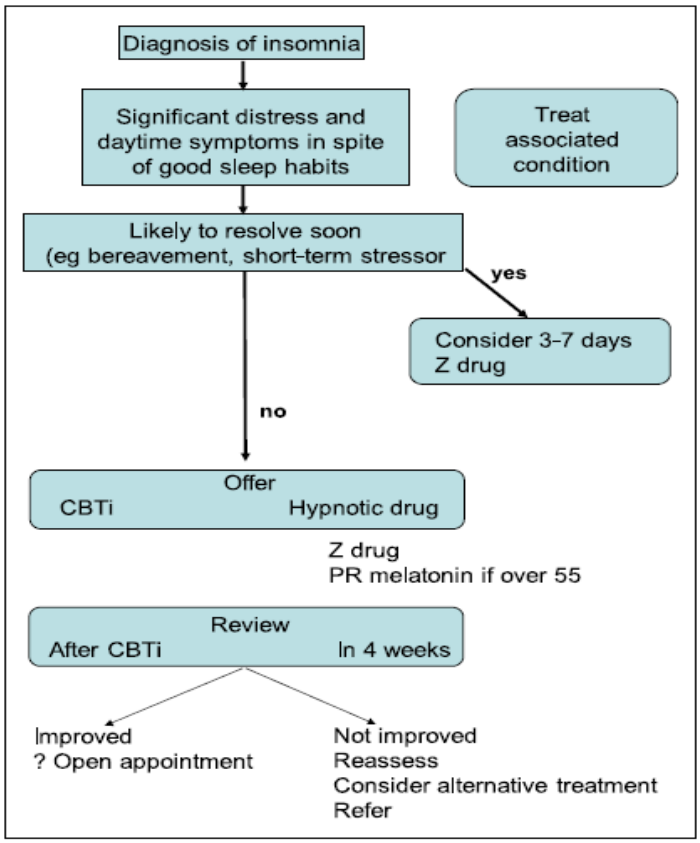
HCP = Health care professionals; OTC = over-the-counter; Rx = Medical prescription.

Clinical treatment pathway

Given the chronicity of insomnia disorder, longer-term treatment is anticipated as part of its clinical treatment pathway. However, none of the existing guidelines include recommendations for long-term treatment of insomnia disorder. The pathways presented below highlight short-term treatment options that are often used for longer than recommended, (77).

The optimal goal of therapy in insomnia disorder is to improve sleep and daytime functioning (14, 51). Figure 6 illustrates the clinical treatment pathway for insomnia disorder. The NICE Clinical Knowledge Summary (CKS) recommends managing patients with insomnia disorder in primary care for at least 6 months prior to specialist referral, if necessary (51). Monitoring response to treatment is an important element in the management of insomnia disorder, partly driven by the safety concerns associated with current pharmacotherapy. The CKS recommends periodic review (i.e., every 2–4 weeks) of insomnia symptoms to evaluate the need to continue treatment (51).

Figure 6: Insomnia treatment pathway (14)



CBTi = Cognitive behavioural therapy for insomnia; PR= prolonged-release ; Z drug= non-benzodiazepines.

Non-pharmacological treatment

Current guidelines recommend CBTi as the first-line treatment for insomnia disorder (14, 25, 51). Depending on the needs of a patient, CBTi is administered alone or is augmented with short-term pharmacotherapy. CBTi comprises five major components, namely sleep hygiene, sleep education, stimulus control, muscle relaxation, and sleep restriction. It can be delivered by psychologists, primary care practitioners including nurses, or self-administered (e.g., digital applications) (14, 25, 51).

Sleepio®, a self-help digital sleep improvement programme based on CBTi, includes a sleep test, weekly interactive CBTi sessions and regular sleep diary entries. It is designed to be completed in six weeks; however, users have full access to the programme for up to 12 months (74). NICE has recently recommended Sleepio® as a cost-saving treatment option for insomnia in primary care for people who would otherwise be offered sleep hygiene or sleeping pills (74). The availability of Sleepio® is likely to improve the poor access and availability of CBTi in England.

Pharmacotherapy

Pharmacotherapy is considered for treatment naïve patients when CBTi is unavailable, unsuitable or ineffective. The NICE CKS, BAP and European Sleep Research Society (ESRS) guidelines recommend (14, 25, 51):

- BZs and Z-drugs as efficacious hypnotics that should be restricted to the shortest possible duration (<4 weeks) due to safety risks associated with prolonged usage. These safety risks include cognitive impairment, daytime somnolence, tolerance and dependence (78-80). In addition, abrupt discontinuation can lead to rebound insomnia and withdrawal symptoms (81).

- Use of prolonged-release melatonin in patients aged ≥55 years.. The recommended initial duration of treatment is 3 weeks. If adequate response is achieved, the treatment can be continued for another 10 weeks. While melatonin has a better safety profile than BZs and Z-drugs, it is associated with daytime somnolence and should be used with caution in the elderly (82).

The ESRS and BAP guidelines recommend intermittent dosing of BZs and Z-drugs to mitigate safety risks. The guidelines also recommend the use of sedating antidepressants for short-term treatment of insomnia disorder (14, 83). The BAP guidelines do not recommend anti-psychotics as first line pharmacotherapy, while the ESRS guideline recommend against using anti-psychotics for insomnia disorder (14, 25). Non-selective antihistamines have limited role in the treatment of insomnia disorder due to its lack of evidence (14, 25).

None of the currently recommended drug classes are indicated for long-term treatment of insomnia disorder due to safety concerns and the risk of developing tolerance and/or dependence. Yet, these drug classes are commonly used beyond their recommended duration (i.e., <4 weeks for hypnotics, ≤13 weeks for melatonin). A UK insomnia market landscape analysis showed that patients on average were on prescription drugs for [redacted] days in 2021 (62). Specifically, the average duration of therapy was [redacted] days for zopiclone, [redacted] days for melatonin and [redacted] days for amitriptyline (76). A survey among UK-based clinicians showed that only [redacted] perceived current prescription drugs as having a positive impact on insomnia. When

asked about the reasons for prescribing these drugs beyond the recommended duration, [redacted] cited a lack of long-term options as the main reason (77).

B.1.3.6 Unmet need

- Patients with insomnia disorder suffer from impaired QoL and reduced productivity (47).
- Both face-to-face and digital CBTi (e.g., Sleepio®) have [redacted]. Among patients who are eligible for CBTi, [redacted] achieve the desired results (1).
- None of the commonly prescribed insomnia treatments in the UK fulfil the criteria of an ideal treatment.
- There is a need for an evidence-based treatment for insomnia disorder that is safe and effective for longer-term use. This will have an immediate impact on patients' QoL and productivity, which will be important for the post-COVID-19 recovery of the economy.

As discussed in earlier sections, the burden of insomnia disorder is high due to several factors, such as an ageing population, and the increasing prevalence of chronic conditions, including mental health diseases. Approximately 3.3 million adults in England suffer from insomnia disorder, with a substantial impact on patients' QoL and productivity (47). Primary care clinicians are [redacted] (62). Consequently, [redacted], further contributing to the burden of this condition (62).

While CBTi is the recommended first-line treatment for insomnia disorder, poor access and availability of face-to-face CBTi has been a longstanding problem in the UK (74). Being resource intensive, CBTi is administered for a maximum of 6 weeks per patient in the NHS (84). Adherence to CBTi is often poor, with patients having to invest personal time and be disciplined to practise CBTi measures (85). In addition, the lack

of a standardised accredited training for CBTi can lead to inconsistent results (85). Although NICE's recommendation of Sleepio® for treating insomnia is likely to significantly improve access and reduce cost of CBTi (74), [redacted] (1). In addition, sleep restriction, a behavioural component of CBTi, while being an efficient way of treating insomnia disorder (86), is often associated with impaired daytime functioning, interfering with attention and other cognitive processes (87).

An ideal treatment for insomnia disorder would reduce symptoms of sleep disturbance, induce sleep rapidly and maintain it throughout the night, with no alteration of the individual's sleep pattern, maintain its efficacy over the long term with no tachyphylaxis or tolerance, exhibit no next-day residual effects, no risk of dependence or no rebound insomnia, and/or withdrawal symptoms upon treatment discontinuation, and improve daytime functioning. Moreover, the treatment should have an excellent safety profile with minimum side effects, be appropriate for adults and the elderly and be used in the presence of comorbidities or concurrent therapies (8, 14, 88). As shown in Table 3, none of the commonly prescribed insomnia treatments in the UK fulfil the criteria of an ideal treatment (76). Thus, there is a need for a new treatment approach of insomnia disorder that fulfils most, if not all, the characteristics of an ideal treatment.

Table 3: Assessment of commonly prescribed insomnia treatment in the UK based on characteristics of an ideal treatment for insomnia disorder

Characteristics (8, 14, 88)	Nitrazepam, Temazepam (89-92)	Zopiclone (89, 91-93)	Melatonin (94, 95)
Induces sleep rapidly	✓	✓	✓
Maintains sleep throughout the night	✓	✓	x
Preserves sleep architecture	x	x	✓
Improves daytime functioning	x	x	x
Indicated for long term use	x	x	x
No next-day residual effects	x	x	x
No risk of dependence/ tolerance	x	x	✓
No rebound insomnia / withdrawal upon discontinuation	x	x	✓
Appropriate for adults and elderly	x	x	x
Minimal important interactions	x	x	✓

UK = United Kingdom.
Note: Fulfilment of each characteristic is based on non-comparative evidence; therefore a direct comparison should not be made between the drugs.

The burden of insomnia disorder remains high despite the existing treatment armamentarium.

Positioning of daridorexant in the current clinical pathway

Based on the population studied in the phase III confirmatory and extension trials of daridorexant (study 301 and 303) and considering NICE's recommendation for Sleepio®, the optimal positioning of daridorexant for the treatment of insomnia disorder in primary care would be:

1. For treatment experienced patients who have already completed standard of care including pharmacotherapy, daridorexant can be an alternative option.

2. For treatment-naïve patients who fail to respond to digital or face-to-face CBTi, daridorexant may be administered as a second-line treatment.

3. Where digital or face-to-face CBTi is inaccessible, or where a patient is unable to follow CBTi steps, or refuses CBTi, daridorexant may be administered as an alternative first-line treatment.

4. When longer-term management of insomnia symptoms (i.e., beyond 4 weeks) is required, daridorexant may be administered as maintenance treatment.

5. When a patient is awaiting access to CBTi or a sleep specialist, daridorexant may be administered to provide rapid symptom relief.

Daridorexant is the first DORA to be approved in the UK and Europe for the treatment of insomnia disorder. It is an evidence-based treatment with established efficacy and safety for up to one year. The clinical trials of daridorexant demonstrated improved night-time and daytime symptoms of insomnia in adult and elderly patients with a favourable safety profile. Patients who received daridorexant did not experience next-morning residual symptoms or show signs of abuse at the recommended dosages of 25 and 50 mg daily. In addition, patients treated with daridorexant did not show any signs of rebound insomnia or withdrawal symptoms upon treatment discontinuation (96, 97).

A NICE recommendation for daridorexant can provide access to an efficacious, safe, and long-term treatment alternative for patients with insomnia disorder. Early treatment with daridorexant can help improve patients' QoL, thereby improving their work productivity and minimising exposure to off-label and potentially harmful pharmacotherapies. Given the high refusal and failure rates of CBTi, a broad recommendation for daridorexant will add an efficacious and safe treatment to the insomnia disorder treatment armamentarium in primary care. This is consistent with the opinion of clinicians who highlighted that the realistic chances of expansion of current services or having trained resources to meet the demands of insomnia patients is minuscule (98). Hence, they consider daridorexant to be a suitable treatment that can be made available for prescription in primary care, especially for long-term use considering the pervasive nature of insomnia disorder (98). Furthermore, since GPs are already experienced in prescribing for this patient group, they can introduce daridorexant within the current primary care pathways without any additional burden on the system.

B.1.4 Equality considerations

While guidelines recommended CBTi as first-line treatment for insomnia disorder, [redacted] of patients refuse CBTi when recommended by their GPs. Among those who receive either face-to-face or digital CBTi, [redacted] fail to achieve the desired results, leading to an overall CBTi success rate of only [redacted] (1). Therefore, a broad recommendation for daridorexant to treat insomnia disorder in primary care will provide GPs with a safe and effective option for patients who refuse or fail CBTi.

B.2 Clinical effectiveness

- Daridorexant 50 mg demonstrated superior efficacy and comparable safety to placebo in patients with insomnia disorder.
- This company submission presents the clinical effectiveness and safety of daridorexant 50 mg in confirmatory study 301 and extension study 303, conducted among adult and elderly patients with insomnia disorder.
- Confirmatory study 301 was a double-blind randomized controlled trial (RCT) which enrolled 930 subjects with insomnia disorder, randomly assigned to receive daridorexant 50 mg (N=310) or placebo (N=310) for 12 weeks.
- Extension study 303 was primarily a comparative safety study, but it included placebo-controlled subjective efficacy data of relevance to assess long-term maintenance with daridorexant. Subjects assigned to daridorexant 50 mg in confirmatory study 301 continued the same dose in study 303 (N=137), while those assigned to placebo were re-randomized to receive either placebo (N=128) or daridorexant 25 mg. The treatment period lasted 40 weeks (for a total of 52 weeks for study 301 and 303).
- In study 301, reductions from baseline in wake after sleep onset (WASO) and latency to persistent sleep (LPS) were greater for daridorexant 50 mg compared to placebo at both month 1 (least squares mean [LSM] difference -22.78 minutes [min] 95% confidence limit [CL] -28.00 to -17.57], p<0.0001 for WASO; -11.35 min [-16.02 to -6.69], p<0.0001 for LPS) and month 3 (-18.30 min [-23.95 to -12.66], p<0.0001 for WASO; -11.67 min [-16.35 to -6.99], p<0.0001 for LPS).
- In terms of key secondary endpoints, daridorexant 50 mg demonstrated:
    - Significant improvement in subjective total sleep time (sTST) compared to placebo at month 1 (LSM difference 22.06 min [14.41 to 29.71], p<0.0001) and month 3 (19.77 min [10.62 to 28.92], p<0.0001).

- Significant improvement in IDSIQ sleepiness domain score compared to placebo at month 1 (-1.75 [-2.51 to -0.98], p<0.0001) and month 3 (-1.90 [- 2.90 to -0.91], p=0.0002).

- For other secondary endpoints, daridorexant 50 mg was superior to placebo*:
  - It led to [redacted] in total sleep time (TST) at month 1 [redacted] and month 3 [redacted].
  - It led to [redacted] in subjective WASO (sWASO) at month 1 [redacted] and month 3 [redacted].
  - Similarly, it led to [redacted] in subjective latency to sleep onset (sLSO) at month 1 [redacted] and month 3 [redacted].

- Subgroup analyses of primary and key secondary efficacy endpoints revealed that daridorexant 50 mg was consistently superior to placebo across all pre-specified subgroups of age, sex, and region.

- The superiority of daridorexant 50 mg compared with placebo for both objective and subjective measures of insomnia were supported by extension study 303. This indicated that the benefits were sustained for up to a year on treatment*:
  - sTST [redacted] from confirmatory baseline to month 6 compared with placebo [redacted], but not in months 9 [redacted] and 12 [redacted].
  - IDSIQ sleepiness domain score [redacted] from confirmatory baseline compared with placebo at month 6 [redacted], month 9

[redacted] and month 12 [redacted].

- Additional analysis of ISI© scores showed that*:
  - Daridorexant 50 mg [redacted] ISI© scores from baseline compared with placebo at month 1 [redacted] and at month 3 [redacted].
  - ISI© scores [redacted] over time indicating [redacted] in both treatment groups of extension study 303, attributed to selective attrition.
- Daridorexant 50 mg demonstrated a favourable safety profile in studies 301 and 303*. Study subjects who received daridorexant did not experience withdrawal symptoms upon treatment discontinuation. Further, daridorexant use indicated no signs of impaired daytime functioning.

*Endpoints not statistically powered to demonstrate significance.

B.2.1 Identification and selection of relevant studies

A SLR was conducted to identify RCT evidence on the efficacy and safety of pharmacological treatments for subjects with insomnia disorder from the published literature. Searches were conducted in literature databases, clinical trial registries and conference proceedings. All searches were last conducted on 1st March 2022. The eligible studies encompassed all RCTs evaluating the efficacy of pharmacological interventions used in the treatment of adults (age ≥18 years) with insomnia disorder (Appendix D).

Study 301 (NCT03545191), the pivotal phase III study and its safety and tolerability extension study 303 (NCT03679884) were identified as the relevant clinical studies for this appraisal (Appendix D). Both clinical trials assessed the 50 mg once daily dosage of daridorexant compared with placebo – based on which they were chosen to be reviewed for this appraisal.

Full details of the search strategy, identification, selection and synthesis of clinical evidence relevant to the technology being appraised are described in Appendix D.

B.2.2 List of relevant clinical effectiveness evidence

The daridorexant clinical trial programme for insomnia disorder encompasses five trials (i.e., two phase II studies [201 and 202], two phase III confirmatory studies [301 and 302] and one phase III extension study [303, Table 4]). Study 301 assessed 25 mg and 50 mg doses of daridorexant. Study 303 assessed 10 mg, 25 mg and 50 mg doses of daridorexant. For this company submission, only confirmatory study 301 and the extension study 303 are considered relevant to the NICE scope and the evidence for daridorexant 50 mg versus placebo is presented. The other phase III study 302 was excluded as it evaluated the 10 mg and 25 mg doses of daridorexant, which is not relevant for this appraisal.

Evidence from confirmatory study 301 and extension study 303 are presented in this company submission to support the NICE appraisal of the 50 mg dose of daridorexant for the population of interest and serve as the primary source of clinical effectiveness data for daridorexant in the cost-effectiveness analysis described in B.3 Cost effectiveness. The clinical effectiveness evidence available from studies 301 and 303 are summarised in Table 5 and Table 6.

Table 4: Overview of studies in the clinical trial programme of daridorexant

Study name	Study identifier	Main objective	Status	Relevant for this appraisal & reason
201	NCT02839200	Assess the efficacy and safety of daridorexant in adult subjects with insomnia disorder (dose response study)	Completed	No, it was a dose finding study
202	NCT02841709	Assess the efficacy and safety of daridorexant in elderly subjects with insomnia disorder	Completed	No, it was a dose finding study
301	NCT03545191	Assess the efficacy and safety of daridorexant in adult and elderly subjects with insomnia disorder	Completed	Yes, meets the PICO criteria as defined in the decision problem
302	NCT03575104	Assess the efficacy and safety of daridorexant in adult and elderly subjects with insomnia disorder	Completed	No, doses of daridorexant used are not relevant for the decision problem
303	NCT03679884	Assess the long-term safety and tolerability of daridorexant in adult and elderly subjects with insomnia disorder	Completed	Yes, meets the PICO criteria as defined in the decision problem

PICO=population, intervention, comparator and outcome.

Table 5: Clinical effectiveness evidence for study 301 (99)

Study	ID-078A301 (NCT03545191)					
Study design	Multi-centre, double-blind, randomized, placebo-controlled, parallel-group					
Population	Adult (18-64 years) and elderly (≥ 65 years) male and female subjects with a diagnosis insomnia disorder as per the DSM-5® criteria and moderate-to-severe insomnia as per ISI© (ISI© ≥ 15).					
Intervention(s)	Daridorexant (25 mg and 50 mg)*					
Comparator(s)	Placebo					
Indicate if trial supports application for marketing authorisation	Yes	Y	**Indicate if trial used in the economic model**	Yes	Y	
	No			No		
Rationale for use/non-use in the model	Study meets the PICO criteria defined in the decision problem.					
Reported outcomes specified in the decision problem	The outcomes relevant for the decision problem include: 1. Improvement of night-time symptoms of insomnia (WASO, sWASO, LPS) 2. Changes in sleep architecture and sleep efficiency (LPS, TST, sTST) 3. Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function (TST, sWASO, sLSO) 4. Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score 5. Safety and tolerability (adverse events, next morning residual effect, rebound insomnia, abuse potential, SDS©) 6. HRQoL (ISI© score)					
All other reported outcomes	1. Withdrawal symptoms 2. Sleep continuity (WASO by quarter of the night and by hour of the night, TST by quarter of the night, sleep awakenings measured by PSG or self-reported) 3. Sleep efficiency 4. PGA-S, and PGI-C scores					

*Only the evidence for daridorexant 50 mg vs placebo is presented in this submission
DSM®-5=Diagnostic and Statistical Manual of Mental Disorders®, Fifth Edition; IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISI©=Insomnia severity index©; LPS=latency to persistent sleep; PGA-S=Patient Global Assessment of Disease Severity; PGI-C=Patient Global Impression of Change; PICO=population, intervention, comparator and outcome; PSG=polysomnography; REM=rapid eye movement; SDS©= Sheehan disability scale©; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake time after sleep onset; TST= total sleep time; WASO=wake after sleep onset.

Table 6: Clinical effectiveness evidence for study 303 (97)

Study	ID-078A303 (NCT03679884)				
Study design	Multi-centre, double-blind, parallel-group, randomized, placebo-controlled, three doses, 40-week safety extension study to ID-078A301 and ID-078A302				
Population	Adult (18-64 years) and elderly (≥ 65 years) male and female subjects with insomnia disorder according to DSM-5® criteria, who had completed daridorexant treatment in study 301 and 302				
Intervention(s)	Daridorexant (10 mg, 25 mg and 50 mg)*				
Comparator(s)	Placebo				
Indicate if trial supports application for marketing authorisation	Yes	Y	**Indicate if trial used in the economic model**	Yes	Y
	No			No	

Study	ID-078A303 (NCT03679884)
Rationale for use/non-use in the model	Study meets the PICO criteria defined in the decision problem.
Reported outcomes specified in the decision problem	The outcomes of the decision problem include: 1. Safety and tolerability (adverse events, next morning residual effect, rebound insomnia, abuse potential) 2. Improvement of night-time symptoms of insomnia (sWASO) 3. Changes in sleep architecture and sleep efficiency (sTST) 4. Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function (sLSO) 5. Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score 6. HRQoL (ISI© score)
All other reported outcomes	1. SDQ VAS 2. Withdrawal symptoms 3. Self-reported awakenings 4. PGA-S and PGI-C scores

*Only the evidence for daridorexant 50 mg vs placebo is presented in this submission
DSM®-5=Diagnostic and Statistical Manual of Mental Disorders®, Fifth Edition; IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISI©=Insomnia severity index©; PGA-S=Patient Global Assessment of Disease Severity; PGI-C=Patient Global Impression of Change; PICO=population, intervention, comparator and outcome; SDQ= Sleep diary questionnaire; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake after sleep onset; TST= total sleep time; VAS=visual analogue scale; WASO=wake after sleep onset.

Definitions of key sleep-related endpoints used in studies 301 and 303

Objective sleep endpoints of time taken to fall asleep (LPS), number of awakenings or time awake whilst in bed (WASO), and total sleep time (TST) were assessed using PSG at baseline and at months 1 and 3 of treatment. Further, they were also assessed immediately after treatment cessation. Subjective sleep endpoints of sTST, sWASO, transition time from wakefulness to sleep (sLSO) were recorded using SDQ. These subjective endpoints were collected daily during baseline assessments, double-blind treatment period and placebo run-out. In addition, daytime functioning was recorded using IDSIQ daily in the evening during baseline assessments, double-blind treatment period and placebo run-out (97, 99).

B.2.3 Study 301 — summary of trial methodology

B.2.3.1 Study sites

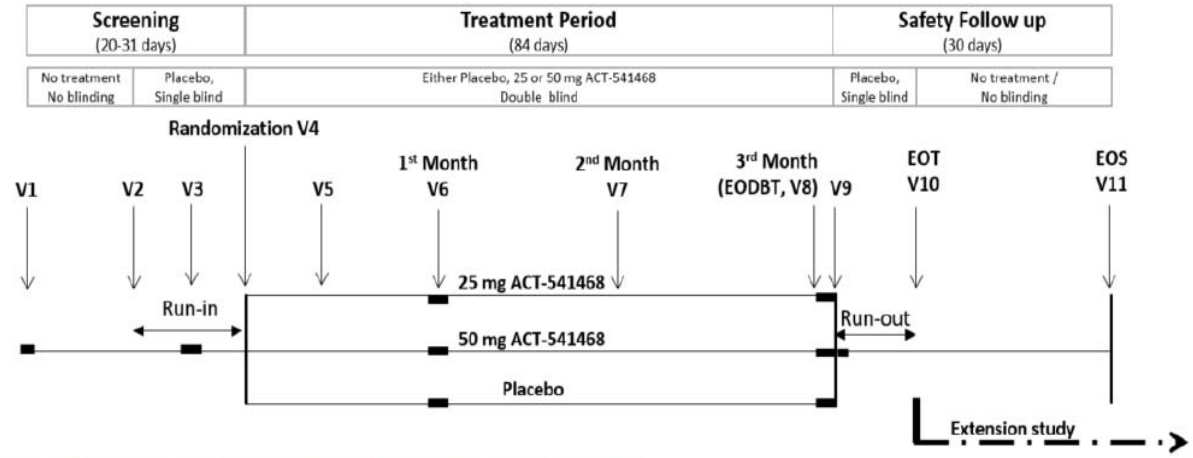
Study 301 involved 75 sites across 10 countries (Australia, Canada, Denmark, Germany, Italy, Poland, Serbia, Spain, Switzerland, and the US), of which 51 sites in seven countries (Canada, Denmark, Germany, Poland, Spain, Switzerland, and the US) enrolled and randomized subjects (99).

B.2.3.2 Study design

Study 301 was a multi-centre, randomized, double-blind, placebo-controlled, parallel-group phase III study evaluating two different doses (25 mg and 50 mg) of daridorexant for 3 months (84 ± 2 days). The study recruited adult and elderly subjects with insomnia disorder, according to the criteria of DSM-5, unless their insomnia was associated with major comorbidities – especially comorbid neurological, affective or psychiatric disorders (e.g., severe or uncontrolled depression or anxiety, dementia) that could interfere with the study endpoints (99).

The overall study design is illustrated in the Figure 7. The core study comprised of three phases – screening, double-blind treatment and safety follow-up (99).

Figure 7: Design of study 301 (99)



V5 and V11 were telephone calls; all other visits were at the site.
— = polysomnography nights; EODBT = end-of-double-blind-treatment; EOS = End-of-Study; EOT = End-of-Treatment; V = Visit.

- **Screening phase:** from signing informed consent at Visit 1 until randomization (Visit 4), lasting 20 to 31 days. Eligibility for the study was assessed at multiple time points during the screening phase, according to the inclusion and exclusion criteria. The screening phase comprised of (99):

    - **Screening period:** from Visit 1 until Visit 2, lasting 7 to 18 days. Following signing of informed consent and initial verification of eligibility, subjects had a one-night PSG assessment (on any night between Visit 1 and Visit 2) and completed a minimum of 7 daily entries in the SDQ.

- **Placebo run-in period:** from Visit 2 until randomization (Visit 4), lasting 13 to 24 days. Following confirmation of eligibility, single-blind placebo treatment once daily (in the evening) was administered during this period from Visit 2 to Visit 4. Visit 3 was considered as the baseline timepoint with two PSG nights performed after subjects had completed at least seven daily SDQ entries (visit 2). The mean of two PSG nights was considered as baseline for the objective endpoints. During the second night of visit 3, subjects self-reported baseline ISI© scores. The mean value based on the screening of SDQ or IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3 was considered as baseline for subjective endpoints.

- **Double-blind treatment phase:** from randomization (Visit 4, Day 1) until end-of-double-blind-treatment ([EODBT] second morning of Visit 8, Day 85). Subjects were randomized in a 1:1:1 ratio to either doses of daridorexant (25 mg or 50 mg), or placebo. Daridorexant was taken by subjects once daily in the evening from Day 1 to Day 84. The SDQ was completed by the subjects daily. A safety telephone call to collect information about adverse events (AEs) and concomitant medications was performed any day from Day 7 to Day 14 (Visit 5). Sleep parameters were objectively assessed twice during two consecutive PSG nights (Days 27 to 29 [Visit 6] and Days 83 to 85 [Visit 8]). A safety visit without PSG night was performed on Day 55 (Visit 7). Subjects who prematurely discontinued study treatment but remained in the study were encouraged to continue with all planned study procedures until end-of-study (EOS), excluding the placebo run-out period.

- **Safety follow-up phase:** from end of double-blind treatment (EODBT) (evening of Day 85, Visit 9) until 30 days after last dose of daridorexant treatment intake for subjects who did not enter study 303, or until enrolment into study 303. This phase comprised of two additional phases (99):

    - **Placebo run-out period:** from the evening on the first day of Visit 9 (Days 85 to 86) until Visit 10 (Day 92), lasting seven days. Subjects received once daily single-blind placebo treatment (Days 85 to 91),and completed the SDQ. Visit 9 consisted of one PSG night (Days 85 to 86). The end of

treatment (EOT) was reached after all visit assessments had been performed at Visit 10.

- **Safety follow-up period:** from EOT (Day 92, Visit 10) until End-of-study (EOS), i.e., the date of enrolment into study 303, or until the 30-day follow-up telephone call (Day 115, Visit 11; collecting information on AEs, serious AEs [SAEs], and concomitant medications) for subjects who did not enter study 303.

For subjects who prematurely discontinued study treatment but did not prematurely withdraw from the study, the follow-up telephone call was performed on Day 115. For subjects who withdrew consent and no longer wished to participate in the study, EOS was the date of consent withdrawal. For subjects declared lost to follow-up, EOS was the date of last successful contact (99).

B.2.3.3 Study eligibility criteria

Table 7 presents the key inclusion and exclusion criteria of confirmatory study 301 (99). The full eligibility criteria are detailed in Appendix M.

Table 7: Inclusion and exclusion criteria of study 301 (99)

Inclusion criteria	• Insomnia disorder according to the DSM-5® criteria. • Self-reported insomnia of at least moderate severity (ISI© score ≥ 15) at screening. • Sleep disturbance causing clinically significant distress or impairment in social, occupational, educational, academic, behavioural, or other important areas of functioning. • Self-reported insufficient sleep quantity (≥ 30 minutes to fall asleep, wake time during sleep ≥ 30 minutes, and sTST ≤ 6.5 hours during the night) for at least 3 nights per week during at least 3 months prior to the screening visit, and for at least 3 out of 7 nights on the SDQ completed during the placebo run-in period prior to the run-in PSG nights. • Objective sleep quantity parameters assessed on 2 consecutive PSG nights during the placebo run-in period: mean LPS ≥ 20 minutes, with neither of the 2 nights < 15 minutes; mean WASO ≥ 30 minutes, with neither of the 2 nights < 20 minutes; and mean TST < 420 minutes. • Subjects were required to sign informed consent prior to any study-mandated procedure.
Exclusion criteria	• Subjects self-reporting daytime napping ≥ 1 hour per day and ≥ 3 days per week. • Subjects with BMI < 18.5 or > 40.0 kg/m^2. • Subjects who were pregnant, breastfeeding, or planning to become pregnant.

	• Subjects with any lifetime history of suicide attempt, sleep-related breathing disorders, periodic limb movement disorder, restless legs syndrome, circadian rhythm disorder, REM behaviour disorder, narcolepsy, or apnoea/hypopnea. • Subjects with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments. • Subjects aged ≥ 50 years with a Mini Mental State Examination© score < 25. • Subjects treated with central nervous system-active drugs; cognitive behavioural therapy was allowed if started at least 1 month prior to the run-in PSG nights and intended to be continued throughout the study. • Subjects not able or willing to stop treatment with moderate or strong CYP3A4 inhibitors or inducers within at least 1 week prior to the start of the placebo run-in period.

BMI=Body mass index; CYP3A4=cytochrome P450 3A4; DSM®-5=Diagnostic and Statistical Manual of Mental Disorders®, Fifth Edition; ISI©=Insomnia severity index©; LPS=Latency to persistent sleep; PSG=Polysomnography; REM=rapid eye movement; SDQ= Sleep diary questionnaire; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake after sleep onset; TST= total sleep time; VAS=visual analogue scale; WASO=wake after sleep onset.

B.2.3.4 Study treatment, prior and concomitant medications

Study treatment comprised of single-blind treatment (placebo matching daridorexant, administered during the placebo run-in and run-out periods) and double-blind treatment (daridorexant, or placebo matching daridorexant, administered from randomization to EODBT) (Table 8) (99).

Table 8: Trial drugs in study 301 (99)

Drug	Dose	Frequency of administration	Route of administration	Duration
Daridorexant, film coated tablet	25 mg and 50 mg	One tablet taken orally once daily in the evening	Oral	84 ± 2 days
Placebo matching daridorexant, film coated tablet	-			Single-blind placebo run-in period (13–24 days), treatment period (84 ± 2 days), and Single-blind placebo run-out period (7 + 2 days)

CBTi was only allowed if the treatment started at least one month prior to Visit 3 (baseline) and the subject agreed to continue this CBTi throughout the study. Initiation of CBTi during the study was not allowed (99).

Therapies considered necessary for a subject's well-being and not categorized as prohibited concomitant medications could be used in this study. However, initiation of

new medication was discouraged, and concomitant medication was preferably not changed during the study (99).

The use of non-sedating antihistamines, opioids/narcotics, centrally acting muscle relaxants with psychotropic effects, and pseudoephedrine was permitted with restrictions. Inhaled or nasal corticosteroids were permitted (99).

The following concomitant therapies were forbidden during the study (99):

- Treatment with another investigational drug until EOS.
- Study-prohibited central nervous system (CNS)-active medications for 5 half-lives of the respective drug (but at least 2 weeks) prior to Visit 1 and until 24 h after EOT.
- Treatment with moderate or strong CYP3A4 inhibitors, or moderate or strong CYP3A4 inducers until 24 hours after EOT.

B.2.3.5 Pre-specified study endpoints

The pre-specified endpoints relevant for the decision problem are summarised in Table 9 (99).

Table 9: Primary, secondary and exploratory endpoints of study 301 relevant to NICE decision problem (99)

Primary efficacy endpoints	**Definition**	**NICE scope/ economic model?**
Change from baseline[a] to month 1[b] and 3[b] in WASO	WASO was the time spent awake after onset of persistent sleep (beginning of the first continuous 20 epochs [i.e., 10 min] scored as non-awake, i.e., epochs scored as either S1, S2, SWS or REM until lights on, as determined by PSG.	Per NICE scope, not included in the economic model
Change from baseline[a] to month 1[b] and 3[b] in LPS	LPS was the time from start of recording to the beginning of the first continuous 20 epochs (i.e., 10 min) scored as non-awake, i.e., epochs scored as either S1, S2, SWS or REM, as determined by PSG.	
Secondary efficacy endpoints	**Definition**	**NICE scope/ economic model?**
Change from baseline[c] to month 1[d] and 3[d] in sTST	sTST was the time reported by the subject in answer to the SDQ question “In total, how long did you sleep last night? (This should just be your best estimate, based on when you went to bed and woke up, how long it took you to fall asleep, and how long you were awake. You do not need to	Per NICE scope, not included in the economic model

	calculate this by adding and subtracting; just give your best estimate.)" Nightly sTST were averaged over 7 nights preceding the visit at end of Month 1 and Month 3	
Change from baseline[c] to month 1[d] and 3[d] in IDSIQ sleepiness domain score	The IDSIQ, a patient-reported outcome measure of subjects' perception of their daytime symptoms of insomnia has a total score ranging from 0 to 140. IDSIQ sleepiness domain score, based on the subject's responses for 4 items, could range from 0 to 40 (whole numbers only) with higher scores indicating greater burden of illness during the daytime.	
Other efficacy endpoints	**Definition**	**NICE scope/ economic model?**
Change from baseline[a] to month 1[b] and 3[b] in TST	TST was the time scored as non-awake (i.e., S1, S2, SWS, or REM) from lights off to lights on, as determined by PSG.	Per NICE scope, not included in the economic model
Change from baseline[c] to month 1[d] and 3[d] in sWASO	sWASO was the time spent awake after sleep onset reported by the subject in answer to the SDQ question "In total, how long did these awakenings last?"	
Change from baseline[c] to month 1[d] and 3[d] in sLSO	sLSO was the time reported by the subject in answer to the SDQ question "How long did it take you to fall asleep?"	
Change from baseline[c] to month 1[d] and 3[d] in IDSIQ scores	The IDSIQ total score is the sum of the three IDSIQ domain scores: alert/cognition (0 to 60), mood (0 to 40) and sleepiness (0 to 40). Higher scores indicate greater burden of illness during the daytime.	
Exploratory endpoints	**Definition**	**NICE scope/ economic model?**
Other exploratory endpoints	*The statistical methods and results of ISI© are reported in B.2.3.6 and B.2.4.5, respectively.* • Change from baseline (Visit 3) to Month 1 (Visit 6) and Month 3 (Visit 8) in ISI© scores	Per NICE scope, included in the economic model
	The statistical methods and results of the following exploratory endpoints are reported in Appendix M. • Change from baseline to Month 1 and Month 3 in WASO over time (by hour of the night and by quarter of the night)[e] • Change from baseline to Month 1 and Month 3 in sleep quality, depth of sleep, daytime alertness, and daily ability to function, as determined by scores on the VAS[f]	Per NICE scope, not included in the economic model

	• Change from baseline to Month 1 and Month 3 in duration of TST in each sleep stage (S1, S2, SWS and REM)[e] • Change from baseline to Month 1 and Month 3 in sleep architecture assessed as percentage of TST in each sleep stage (S1, S2, SWS, and REM) over the whole night, and for each quarter of the night[e] • Change from baseline to Month 1 and Month 3 in mean numbers of shifts from S2, SWS or REM to S1 or awake[e] • Change from baseline to Month 1 and Month 3 in mean number of awakenings (defined as the number of awakenings between first epoch and last epoch not scored wake) as measured by PSG (for the whole night, by quarter of the night, and by hour of the night)[e] • Change from baseline to Month 1 and Month 3 in mean number of self-reported awakenings[f] • Change from baseline to Month 1 and Month 3 in sleep efficiency (defined as 100 [TST/time in bed])[e] • Change from baseline to Month 1 and Month 3 in Patient Global Assessment of Disease Severity (PGA-S scores [daytime symptoms])[g] • Change from baseline to Month 1 and Month 3 in Patient Global Impression of Change (PGI-C scores [daytime symptoms])[g] • Change from baseline to Month 1 and Month 3 in PGI-C scores (night-time symptoms)[g] • Change from baseline to Month 1 and Month 3 in PGI-S scores (night-time symptoms)[g]	
Safety	*The statistical methods and results of the following safety endpoints are reported in B.2.3.6 and B.2.5, respectively.* • TEAEs up to 30 days after double-blind study treatment discontinuation or until enrolment into study 303 • SAEs up to 30 days after double-blind study treatment discontinuation or until enrolment in the extension study • AEs leading to premature discontinuation of treatment • AESI after adjudication by the ISB (narcolepsy-like symptoms or suicide/self-injury) • Withdrawal effects (physical dependence) upon treatment discontinuation, assessed based on the changes in the BWSQ total score from last assessment on double-blind treatment (Visit 8, 2nd morning) and the placebo run-out period (Visit 9 and Visit 10), the occurrence of relevant AEs, and marked ECG abnormalities. The BWSQ consists of 20 items and is used to assess the main symptoms which might be experienced during withdrawal from benzodiazepines. Symptoms are rated as 0 (No), 1 (Yes-moderate), or 2 (Yes-severe).	Per NICE scope, not included in the economic model

	• Changes from baseline to Month 1 and Month 3 in SDS© [i] [questionnaire on impairment of work, social life, and family life/home responsibilities, each on a 10-point scale]	
	The results of additional safety endpoints are reported in Appendix M. • Change from baseline (Visit 3) to Month 1 (Visit 6) and Month 3 (Visit 8) in vital signs (mean of the 2 PSG nights in systolic and diastolic BP and pulse rate) • Change from baseline (Visit 1) to Month 3 (Visit 8) in body weight. • Marked ECG abnormalities on double-blind study treatment. • Change from baseline (Visit 3) to Month 3 (Visit 8) and the end of the placebo run-out period (Visit 10) in ECG parameters. • Marked laboratory abnormalities on double-blind study treatment. • Change from baseline (Visit 3) to Month 1 (Visit 6) and Month 3 (Visit 8) in laboratory parameters. • Occurrence of suicidal ideation and/or behaviour on double-blind study treatment based on C-SSRS© (the presence and severity of both suicidal ideation and behaviours). • Rebound insomnia, assessed based on objective sleep parameters WASO and LPS at the start of the placebo run-out period (Visit 9) as compared to baseline (Visit 3), and on the subjective sleep parameter sTST[h] during the placebo run-out period as compared to baseline. • Next-morning residual effect assessed based on changes from baseline to Month 1 and Month 3 in: o Coding sub-test©[i] [used to measure attention, perceptual speed, motor speed, visual scanning, and memory] o Morning sleepiness score on the VAS[h]	Per NICE scope, not included in the economic model

[a]Baseline: mean of the 2 PSG nights at Visit 3.
[b]Month 1 and Month 3: mean of the 2 PSG nights at Visit 6 and Visit 8, respectively
[c]Baseline: mean value based on the screening SDQ / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3.
[d]Month 1 and Month 3: mean value based on the SDQ / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 6 and Visit 8, respectively.
[e]Baseline: mean of the 2 PSG nights at Visit 3. Month 1 and Month 3: mean of the 2 PSG nights at Visit 6 and Visit 8, respectively.
[f]Baseline: mean value based on the screening SDQ/IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3. Month 1 and Month 3: mean value based on the SDQ/IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 6 and Visit 8, respectively.
[g]Baseline: Visit 3. Month 1: Visit 6 or, if that is missing, week 4 of the questionnaire. Month 3: Visit 8 or, if that is missing, week 12 of the questionnaire.
[h]Baseline: mean value based on the screening SDQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3. Month 1 and Month 3: mean value based on the SDQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 6 and Visit 8, respectively. Placebo run-out period: mean value based on the SDQ entries performed in the 7 days immediately after the PSG night at Visit 9.
[i]Baseline: mean of the two PSG morning assessments at Visit 3. Month 1 and Month 3: mean of the two PSG morning assessments at Visit 6 and Visit 8, respectively.
AEs=Adverse events; AESI= Adverse event of special interest; BWSQ= Benzodiazepine Withdrawal Symptom Questionnaire; C-SSRS©= Columbia Suicide Severity Rating Scale©; ECG= electrocardiogram; IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISB=independent safety board; ISI©=Insomnia severity index©; LPS=latency to persistent sleep; PGA-S=Patient Global Assessment of Disease Severity; PGI-C=Patient Global Impression of Change; PGI-S=Patient Global

Impression of Severity PICO=population, intervention, comparator and outcome; PSG=polysomnography; REM=rapid eye movement; S1, S2, S3= sleep stage 1, 2 and 3; SAE=Serious AEs; SDS= Sheehan Disability Scale©; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake after sleep onset; SWS= slow-wave sleep; TEAEs= Treatment-emergent AEs; TST= total sleep time; VAS=visual analogue scale; WASO=wake after sleep onset.

B.2.3.6 Statistical methods and analysis sets

Table 10 details the statistical methods and analysis sets used in confirmatory study 301 (99).

Table 10: Summary of statistical methods and analysis sets of study 301 (99)

Study name (number)	Study 301 (NCT03545191)
Research hypothesis relevant to NICE scope	For each of the primary endpoints (change from baseline in WASO [sleep maintenance] and LPS [sleep onset], and secondary endpoints (change from baseline in sTST [sleep quantity], and IDSIQ sleepiness domain score [daytime function], four null hypotheses were defined as follows: • H1: Daridorexant 50mg – Placebo = 0 at Month 1 • H2: Daridorexant 50mg – Placebo = 0 at Month 3 where 'Daridorexant 50mg', and 'Placebo' represent the mean change from baseline for the given endpoint (WASO, LPS, sTST or IDSIQ sleepiness domain score) and time point (Month 1 or Month 3).
Analysis sets	• **Screened analysis set:** The Screened analysis set comprised all subjects who entered screening and received a subject identification number. • **Full analysis set:** The Full analysis set comprised all subjects assigned (i.e., randomized) to a double-blind study treatment. In order to adhere to the intention-to-treat principle as much as possible: • **Per-protocol set:** The Per-protocol set comprised all subjects from the Full analysis set who received at least one dose of double-blind study treatment and who complied with the protocol sufficiently to be likely to exhibit the treatment effects. • **Safety set:** The Safety set comprised all subjects who received at least one dose of double-blind study treatment. • **Treatment withdrawal set:** The Treatment withdrawal set comprised all subjects in the Safety set who received at least one dose of single- blind placebo treatment in the placebo run-out period.
Statistical analysis for primary and key secondary efficacy endpoints	Analysis of the primary and secondary efficacy endpoints was performed on the Full analysis set. • Linear mixed effects model was used for the analysis of change from baseline in WASO, LPS, sTST and IDSIQ sleepiness domain score, separately. • The analysis model adjusted for the baseline value of the relevant response variable (either WASO, LPS, sTST or IDSIQ sleepiness domain score), age group (< 65; ≥ 65 years), treatment (daridorexant 50 mg; placebo), visit (Month 1; Month 3), and the interaction of treatment by visit, and baseline by visit. • To evaluate the efficacy hypotheses, appropriate contrasts were computed to test the treatment differences of interest (i.e., the difference

Study name (number)	Study 301 (NCT03545191)
	in LSM change from baseline between daridorexant and placebo, both at Month 1 and Month 3).
Statistical analysis for other efficacy endpoints	Analysis of the other efficacy endpoints was performed on the Full analysis set. • The same model as for the main analysis of the primary and secondary endpoints (linear mixed effects model), was fitted for TST, sWASO, sLSO and IDSIQ scores (total score; alert/cognition and mood domain scores). The LSM for each treatment group was reported with associated SEs and 95% CIs. The placebo-adjusted LSM was displayed with associated SE, 95% CI and unadjusted two-sided p-value. • Other efficacy endpoints (change from baseline to Month 1 and Month 3 in TST, sWASO, sLSO, and IDSIQ total, alert/cognition domain, and mood domain scores), with their observed values, were summarized descriptively.
Statistical analysis of exploratory endpoints	Analysis of the exploratory efficacy endpoints was performed on the Full analysis set. The exploratory endpoints (change from baseline to Month 1 and Month 3 of the respective variables) were summarized descriptively with the observed values.
Statistical analysis of safety endpoints	All safety endpoints were summarised descriptively.
Sample size & power calculation	The assumptions for the between-subject SD per treatment group for WASO, LPS, and sTST were based on the two phase II studies (201 and 202) conducted in adult and elderly subjects with insomnia receiving 5 mg, 10 mg, 25 mg, 50 mg daridorexant or placebo. The difference compared to placebo in the mean change from baseline to Month 1 and Month 3 was assumed to be 15 (WASO and LPS) and 20 minutes (sTST). Based on a two-sample z-test, at least 900 subjects randomized to 50 mg daridorexant, 25 mg daridorexant, and placebo in a 1:1:1 ratio (i.e., 300 per group) would provide 98.9% power to detect an effect size of 0.37 for a single hypothesis test. This accounts for the Bonferroni correction, where the significance level (alpha) is halved and set to 2.5% two-sided. However, as the number of null hypotheses (endpoints) to test increases, the power decreases. The power calculation assumed all null hypotheses were independent (a conservative assumption for power calculations). Consequently, 900 subjects provided at least 90% power to detect an effect size of 0.37 for testing nine independent null hypotheses.

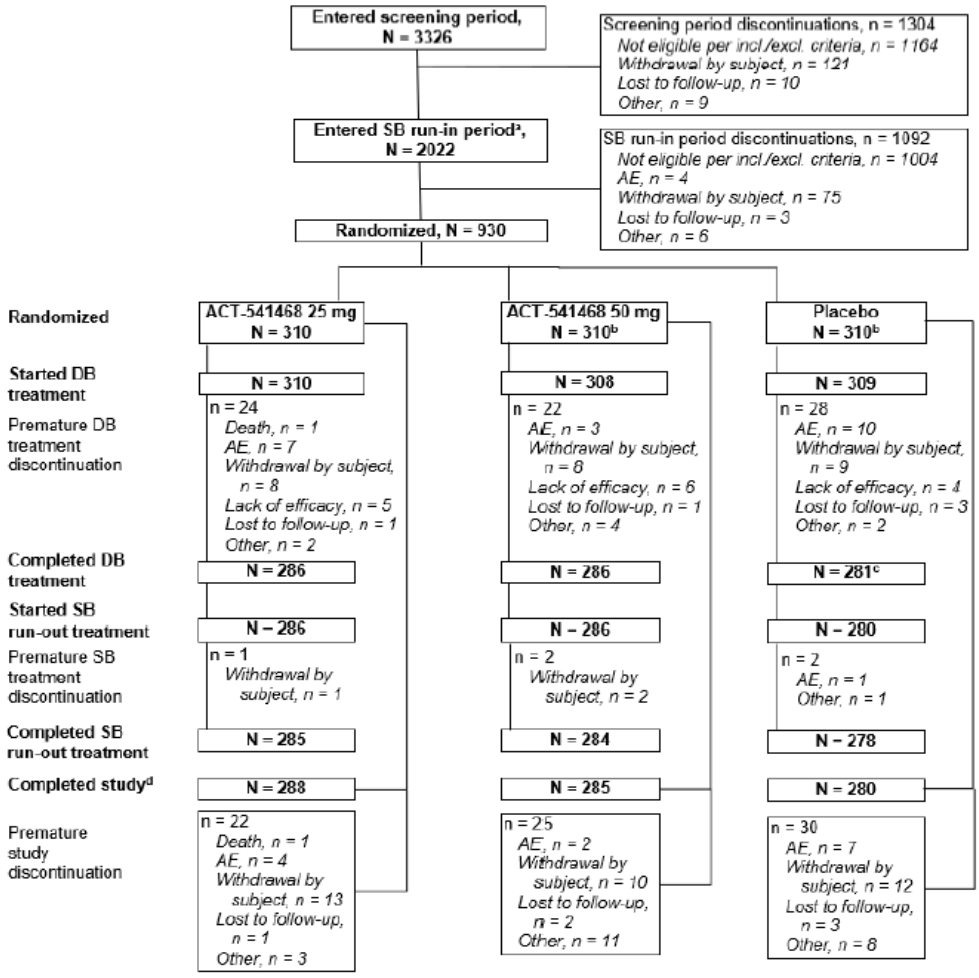
Study name (number)	Study 301 (NCT03545191)
Data management, patient withdrawals	**Handling of partially missing data:** • Partially missing data for WASO and LPS values were handled as follows: if one of the two values was missing either for baseline, Month 1 or Month 3, the single value available was used as the mean for that time point. If both values were missing for a time point, then the mean value was considered missing for that time point. The same approach was used for the following variables: TST, number of shifts from S2, SWS or REM to S1 or awake, number of awakenings, Coding sub-test©, SDS©, and neurological examination. • For sTST and IDSIQ sleepiness domain scores, subjects had to have at least 2 days of data during each week to calculate a weekly mean. Otherwise, the mean value was considered missing for that week. The same approach was used for the following variables: sWASO, sLSO, IDSIQ scores (total score, alert/cognition domain and mood domain scores), VAS scores, and number of self-reported awakenings.

IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; LPS=Latency to persistent sleep; LSM=least squares mean; REM=rapid eye movement; S2=sleep stage 2; SD=standard deviation; SDS©=Sheehan disability scale©; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; SWS=slow-wave sleep; sWASO=subjective wake after sleep onset; TST= total sleep time; VAS=visual analogue scale; WASO=wake after sleep onset.

B.2.3.6 Participant flow

A total of 930 subjects were randomised at baseline (included in the Full analysis set), of which 927 (99.7%) were treated with double-blind study treatment (included in the Safety set). Three subjects (2 [daridorexant 50 mg] and 1 [placebo]) who did not meet eligibility criteria and had been randomized in error were discontinued from the study before receiving double-blind treatment (99). Most randomized subjects completed the double-blind study treatment (92%). The treatment dropout rate (8%) over the three-month double-blind period was small and similarly distributed across treatment groups (7.7% [25 mg], 7.1% [50 mg], and 9% [placebo]). Three randomized subjects (0.3%; 1 subject in each treatment group) were being treated with CBTi at screening. Of the 927 subjects (99.7%) not using CBTi at screening, 25 subjects (2.7%; 11, 7, and 7 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported previous treatment failure with CBTi. An overview of the disposition of subjects is shown in Figure 8 (99).

Figure 8: Disposition of subjects in study 301 (99)



[a]Subject received at least one dose of SB run-in treatment.
[b]3 subjects were discontinued from the study before receiving double-blind treatment as they did not meet eligibility criteria and had been randomized in error.
[c]Subject completed double-blind treatment but did not start run-out treatment; the subject completed the study.
[d]Subject completed the 30-day follow-up telephone call.
AE=adverse event; DB=double-blind; SB=single-blind.

B.2.3.7 Baseline characteristics and demographics

Demographic characteristics of the Full analysis set were balanced across treatment groups (Table 11). The majority of subjects were female (64.2%) and White (88.4%). The median age of study subjects at screening was 58 years (range 21–86 years), with elderly subjects (aged ≥ 65 years) comprising 39.0% of the study population (99). Most of the elderly subjects were aged between 65 and 75 years (32.9% of the study population); subjects aged 75 to < 85 years and ≥ 85 years comprised 5.8% and 0.3% of the study population, respectively (99). The mean (SD) body mass index (BMI) was 26.3 (4.4) kg/m^2; more than half of the subjects were above normal weight, being either overweight (BMI 25.0 to ≤ 30.0, 41.3%) or obese (BMI > 30.0, 17.7%) (99).

Table 11: Demographic characteristics of subjects enrolled in study 301 (99)

Variable Statistic	Daridorexant 50 mg N = 310	Placebo N = 310
Age at screening (years)		
Mean (SD)	55.5 (15.3)	55.1 (15.4)
Median (Min, Max)	58 (21, 86)	58 (19, 83)
Sex [n(%)]		
Male	111 (35.8)	100 (32.3)
Female	199 (64.2)	210 (67.7)
Race [n(%)]		
Black or African American	30 (9.7)	28 (9.0)
American Indian or Alaska Native	1 (0.3)	0
Native Hawaiian or other Pacific Islander	1 (0.3)	0
Asian	4 (1.3)	2 (0.6)
White	274 (88.4)	278 (89.7)
Other	0	2 (0.6)
Ethnicity [n(%)]		
Hispanic or Latino	44 (14.2)	51 (16.5)
Not Hispanic or Latino	265 (85.5)	259 (83.5)
Unknown	1 (0.3)	0
BMI (kg/m^2) at screening		
Mean (SD)	26.273 (4.275)	26.428 (4.118)
Region [n(%)]		
US	97 (31.3)	104 (33.5)
Other (non-US)	213 (68.7)	206 (66.5)

BMI=Body mass index; SD=standard deviation; US=United States

Baseline insomnia characteristics at screening are summarised in Table 12. Dissatisfaction with sleep quantity or quality, and sleep disturbance causing significant distress or impairment in daytime functioning were reported by all subjects in the Full analysis set as follows: difficulty maintaining sleep (99.8% of subjects), difficulty initiating sleep (99.7%), and early morning awakening (94.9%) (99).

Time since insomnia diagnosis at randomization was balanced across treatment groups, with a median of 6.6 years for the daridorexant 50 mg group, and 8.2 years for the placebo group. Baseline values for the primary and secondary endpoints, and for ISI© score, were balanced across treatment groups (99).

Table 12: Baseline values for WASO, LPS, sTST, IDSIQ sleepiness domain score, and ISI score (99)

	Daridorexant 50 mg N = 310	Placebo N = 310
WASO (min)		

	Daridorexant 50 mg N = 310	Placebo N = 310
n	309	309
Mean (SD)	95.484 (37.813)	102.511 (40.766)
LPS (min)		
n	309	309
Mean (SD)	63.619 (37.389)	66.535 (39.769)
sTST (min)		
n	309	309
Mean (SD)	313.178 (57.597)	315.886 (53.144)
IDSIQ sleepiness domain score		
n	309	308
Mean (SD)	22.479 (7.207)	22.260 (6.947)
ISI© score		
n	308	309
Mean (SD)	19.3 (4.0)	19.2 (4.0)

Higher IDSIQ sleepiness domain score represents greater burden of illness.
IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISI©= Insomnia Severity Index©; LPS=latency to persistent sleep; SD=standard deviation; sTST=subjective total sleep time; WASO=wake after sleep onset.

B.2.4 Study 301 — clinical effectiveness results

B.2.4.1 Primary efficacy endpoints

WASO was significantly reduced from baseline among subjects in the daridorexant 50 mg group compared to subjects in the placebo group at month 1 (LSM difference - 22.78 minutes [min], [95% CI -28.00 to -17.57], p<0.0001) and month 3 (LSM difference -18.30 min, [-23.95 to -12.66], p<0.0001) (Table 13, Figure 9A) (96).

Similarly, LPS showed a significant reduction from baseline among subjects in the daridorexant 50 mg group compared to subjects in the placebo group at month 1 (LSM difference -11.35 min, [-16.02 to -6.69], p<0.0001) and month 3 (LSM difference - 11.67 min, [-16.35 to -6.99], p<0.0001) (Table 13, Figure 9B) (96).

B.2.4.2 Key secondary efficacy endpoints

Compared with placebo, sTST significantly increased from baseline in the daridorexant 50 mg group at month 1 (LSM difference 22.06 min, [14.405 to 29.708], p<0.0001) and month 3 (LSM difference 19.77 min, [10.623 to 28.918], p<0.0001) (Table 14, Figure 9C) (96).

Subjects in the daridorexant 50 mg group reported significant reduction from baseline in IDSIQ sleepiness domain score compared to placebo at month 1 (LSM difference

- 1.75, [-2.51 to -0.98], p<0.0001) and month 3 (LSM difference -1.90, [-2.95 to -0.98], p=0.0002). (Table 14, Figure 9D) (96).

Table 13: Primary efficacy endpoints – WASO and LPS (99)

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo			
					LSM	SE	95% CL	p-value (two-sided)
Between treatment analysis for change from baseline in WASO (min) to Month 1 and Month 3 Full Analysis Set								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	305	-28.98	1.877	[-32.668, -25.299]	-22.78	2.657	[-27.996, -17.567]	<.0001
Placebo (N=310)	299	-6.20	1.899	[-9.928, -2.475]	-	-	-	-
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	287	-29.41	2.031	[-33.399, -25.427]	-18.30	2.875	[-23.945, -12.661]	<.0001
Placebo (N=310)	283	-11.11	2.049	[-15.131, -7.088]	-	-	-	-
Between treatment analysis for change from baseline in LPS (min) to Month 1 and Month 3 Full Analysis Set								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	305	-31.20	1.684	[-34.506, -27.896]	-11.35	2.378	[-16.022, -6.687]	<.0001
Placebo (N=310)	299	-19.85	1.697	[-23.177, -16.515]	-	-	-	-
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	287	-34.80	1.689	[-38.118, -31.490]	-11.67	2.383	[-16.348, -6.994]	<.0001
Placebo (N=310)	283	-23.13	1.697	[-26.464, -19.803]	-	-	-	-

CL=confidence limit; LPS=latency to persistent sleep; LSM=least squares mean; SE=standard error; WASO=wake after sleep onset.

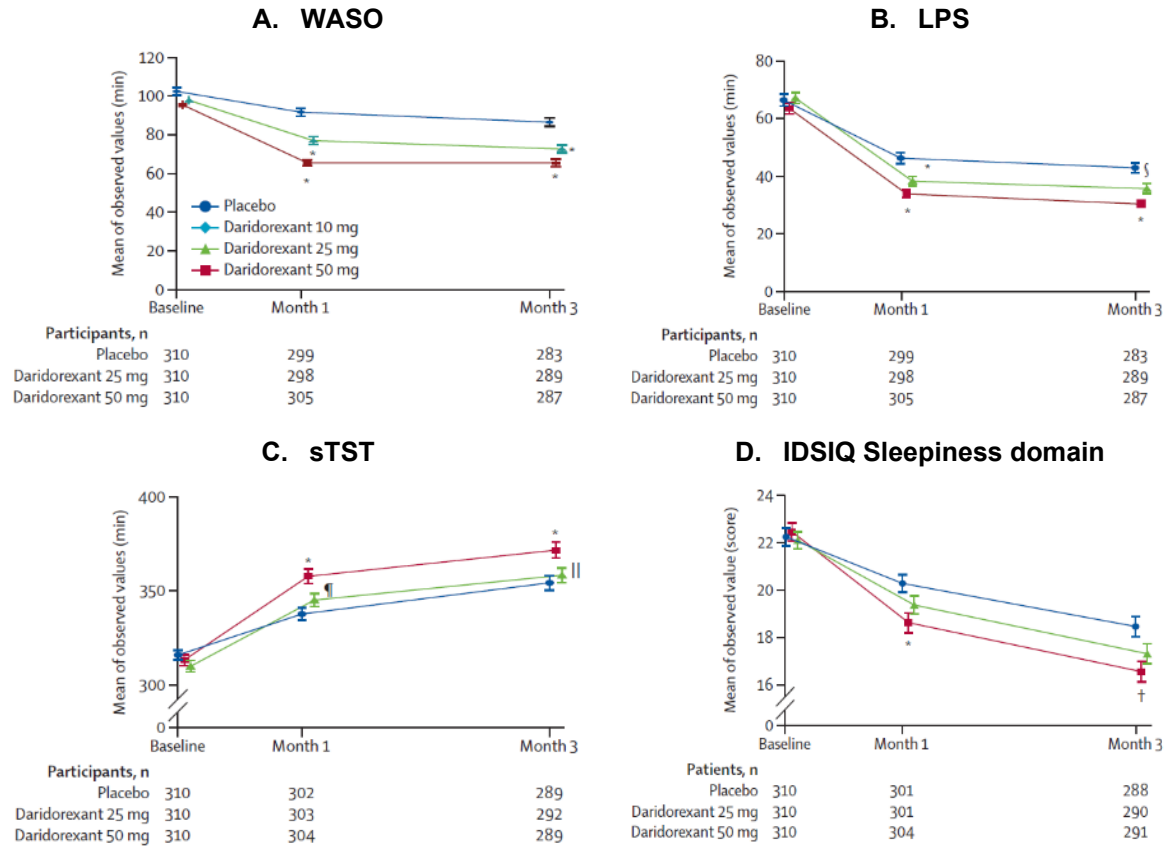
Table 14: Key secondary efficacy endpoints – sTST and IDSIQ sleepiness domain score (99)

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo			
					LSM	SE	95% CL	p-value (two-sided)
Between treatment analysis for change from baseline in sTST (min) to Month 1 and Month 3								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	304	43.62	2.774	[38.173, 49.063]	22.06	3.899	[14.405, 29.708]	<.0001
Placebo (N=310)	302	21.56	2.782	[16.101, 27.022]	-	-	-	-

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo			
					LSM	SE	95% CL	p-value (two-sided)
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	289	57.67	3.311	[51.171, 64.168]	19.77	4.661	[10.623, 28.918]	<.0001
Placebo (N=310)	289	37.90	3.315	[31.393, 44.404]	-	-	-	-
Between treatment analysis for change from baseline in IDSIQ sleepiness domain score to Month 1 and Month 3								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	304	-3.77	0.276	[-4.309, -3.224]	-1.75	0.389	[-2.508, -0.983]	<.0001
Placebo (N=310)	301	-2.02	0.278	[-2.566, -1.476]	-	-	-	-
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	291	-5.70	0.361	[-6.405, -4.987]	-1.90	0.510	[-2.905, -0.905]	0.0002
Placebo (N=310)	288	-3.79	0.363	[-4.503, -3.080]	-	-	-	-

CL=confidence limit; IDSIQ= Insomnia Daytime Symptoms and Impacts Questionnaire; LSM=least squares mean; SE=standard error; sTST= subjective total sleep time.

Figure 9: Night-time efficacy endpoints and IDSIQ sleepiness domain score (96)



Two-sided p-values shown are versus placebo, calculated using the linear mixed effects model for repeated measures. LPS=latency to persistent sleep. sTST=self-reported total sleep time. WASO=wake time after sleep onset. IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire *p<0·0001. †p=0·0001. §p=0·0015. ¶p=0·0013. ||p=0·033.

B.2.4.3 Subgroup analyses of primary and secondary efficacy endpoints

Subgroup analysis was performed to evaluate the consistency of treatment effect across the following demographic subgroups (99):

- Age: < 65, ≥ 65 years
- Sex: Male, female
- Region: US, other (non-US)

The effect of daridorexant 50mg on the primary and key secondary efficacy endpoints was consistent in adults and elderly and across sex and geographical location (Appendix E).

B.2.4.4 Other efficacy endpoints

Daridorexant 50 mg demonstrated significant improvements from baseline across all other efficacy endpoints, including TST, sWASO, sLSO and IDSIQ total, alert/cognition and mood domain scores, compared with placebo at month 1 and month 3 (99).

The analysis of changes from baseline in TST, sWASO, and sLSO is presented in Table 15. Subjects treated with daridorexant 50 mg showed [redacted] in TST compared with placebo at month 1 [redacted] and month 3 [redacted]. sWASO was [redacted] from baseline for daridorexant 50 mg compared with placebo at month 1 [redacted], but not at month 3 [redacted]. In addition, daridorexant 50 mg demonstrated [redacted] from baseline in sLSO compared with placebo at month 1 [redacted] and month 3 [redacted] (99).

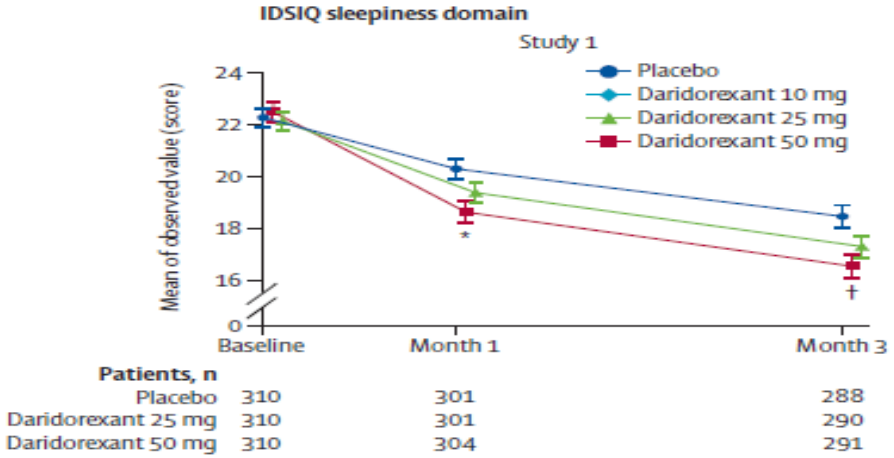
The IDSIQ total (-7.2 [-9.8 to -4.7] at month 1 and -7.2 [-10.5 to -3.9] at month 3), alert/cognition (-2.8 [-3.8 to -1.7] at month 1 and -2.5 [-3.9 to -1.1] at month 3) and mood domain (-2.7 [-3.6 to -1.9] at month 1 and -2.8 [-3.8 to -1.7] at month 3) scores of subjects in the daridorexant 50 mg group showed improvement from baseline compared with placebo at both month 1 and month 3 (all p≤0.001) (Figure 10). Additional analysis of the IDSIQ sleepiness domain score using a 4-point or higher as the meaningful change threshold, yielded numerically higher response rates at both month 1 and month 3 among subjects in the daridorexant 50 mg group compared with the placebo group (127/304 [42%] vs 85/301 [28%] at month 1, 154/291 [53%] vs 128/288 [44%] at month 3) (35, 96).

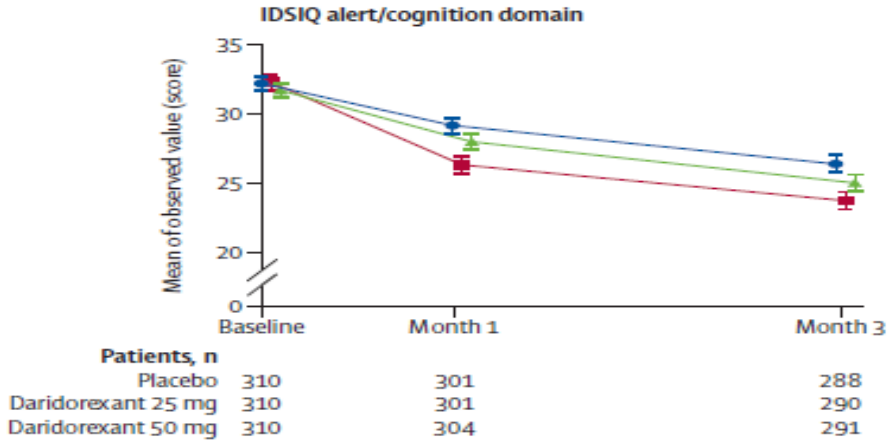
Table 15: Other efficacy endpoints –TST, sWASO, and sLSO (99)

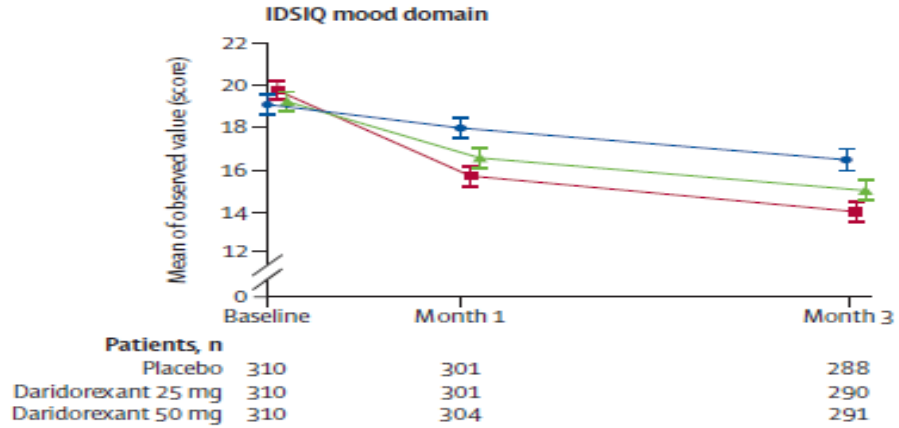
Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo			
					LSM	SE	95% CL	p-value (two-sided)
Between treatment analysis for change from baseline in TST (min) to Month 1 and Month 3								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							
Between treatment analysis for change from baseline in sWASO (min) to Month 1 and Month 3								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							
Between treatment analysis for change from baseline in sLSO (min) to Month 1 and Month 3								
Change from baseline to Month 1								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							
Change from baseline to Month 3								
Daridorexant 50 mg (N=310)	[redacted]							
Placebo (N=310)	[redacted]							

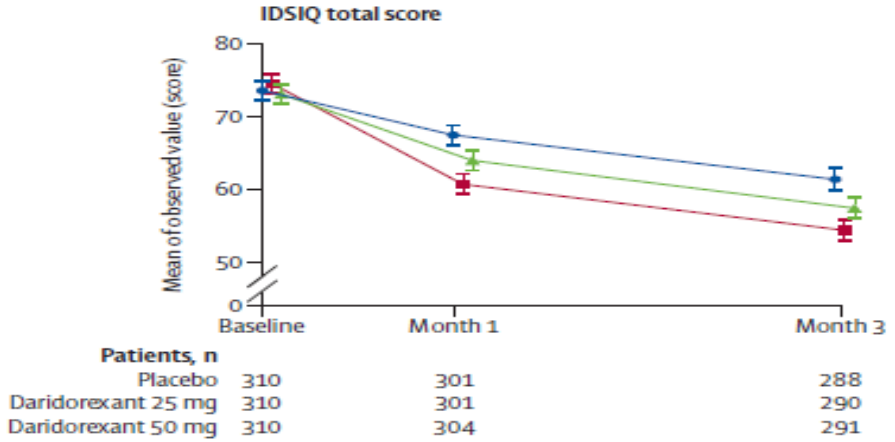
CL=confidence limit; LSM=least squares mean; SE=standard error; sLSO=subjective latency to sleep onset; sWASO=subjective wake after sleep onset; TST=total sleep time.

Figure 10: Other efficacy endpoints – IDSIQ sleepiness domain, IDSIQ alert/cognition domain, IDSIQ mood domain and IDSIQ total score (96)









Two-sided p-values shown are versus placebo, calculated using the linear mixed effects model for repeated measures. p values for the mood domain, alert/cognition domain, and total score comparisons versus placebo (not adjusted for multiplicity). IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire.

B.2.4.5 Exploratory endpoints

As described in B.1.3.1 Disease overview, ISI© scores were used to model the effectiveness of daridorexant compared to placebo and derive EQ-5D utilities for the cost-effectiveness model detailed in B.3 Cost effectiveness (99). The results of the analysis of ISI© scores as a pre-specified exploratory endpoint of confirmatory study 301 are presented below. Additional analysis of ISI© scores necessary for the cost-effectiveness model is presented in B.2.9 Additional analysis of ISI© (99).

Numerically, mean ISI© scores reduced from 19.3 (SD 4.0) at baseline to 14.3 (5.8) at month 1 (mean difference -4.9 [5.5]) and 11.9 (6.3) at month 3 (mean difference -7.2 [6.5]) for daridorexant 50 mg compared with placebo (Table 16) (99). The proportion of subjects who had a decrease in ISI© score of ≥ 6 points from baseline was 40.1% in daridorexant 50 mg compared with 28.6% in placebo at month 1, whereas at month 3 it was 56.5% in daridorexant 50 mg and 46.6% patients in placebo (Table 17).

Results of the other exploratory endpoints supportive of the primary and secondary efficacy endpoints are presented in Appendix M.

Table 16: Exploratory endpoint – ISI© score (99)

Time point Statistic	**n**	**Mean (SD)**
Baseline		
Daridorexant 50 mg (N=310)	308	19.3 (4.0)
Placebo (N=310)	309	19.2 (4.0)
Month 1		
Daridorexant 50 mg (N=310)	299	14.3 (5.8)
Placebo (N=310)	297	16.1 (5.2)
Change from baseline to Month 1		
Daridorexant 50 mg (N=310)	299	–4.9 (5.5)
Placebo (N=310)	297	–3.1 (4.7)
Month 3		
Daridorexant 50 mg (N=310)	283	11.9 (6.3)
Placebo (N=310)	281	13.8 (6.0)
Change from baseline to Month 3		
Daridorexant 50 mg (N=310)	283	–7.2 (6.5)
Placebo (N=310)	281	–5.4 (5.7)

Values for Month 1 / Month 3 were calculated only for subjects who had a baseline value.
A decrease in score represents an improvement.
ISI©=Insomnia Severity Index©;SD=standard deviation.

Table 17: Exploratory endpoint – Subjects with ≥6 points decrease in ISI© score from baseline to month 1 and month 3 (99)

	Daridorexant 50 mg (N = 310) n/Nn (%)	Placebo (N = 310) n/Nn (%)
Month 1 – 2nd Night	120 / 299 (40.1)	85 / 297 (28.6)
Month 3 – 2nd Night	160 / 283 (56.5)	131 / 281 (46.6)

Nn is the number of subjects with non-missing values at the given scheduled visit.
ISI©=Insomnia Severity Index©

B.2.5 Study 301 — adverse reactions

B.2.5.1 TEAEs

During the double-blind study period, 37.7% and 34.0% of subjects reported TEAEs in the daridorexant 50 mg group, and placebo group, respectively. Most of the events were of mild or moderate intensity (99). Nasopharyngitis [6.5% (daridorexant 50 mg); 6.5% (placebo)] and headache [6.2% (daridorexant 50 mg); 3.9% (placebo)] were the most commonly reported TEAEs (Table 18); followed by accidental overdose (2.6% vs 1.6%), fatigue (2.3% vs0.6%), dizziness (2.3% vs 0.6%), and nausea (2.3% vs 1.0%). TEAEs that occurred during the double-blind study period considered related to study treatment were reported for 38 (12.3%), and 29 (9.4%) subjects in the daridorexant 50 mg, and placebo groups, respectively. Fatigue was the most frequent TEAE considered related to study treatment in (1.9% in daridorexant 50 mg vs 0.3% in placebo) (96, 99).

Of note, falls were reported more frequently for placebo (8 subjects) than daridorexant 50 mg group. No suicidal ideation was reported in either treatment groups throughout the entire duration of the study (96, 99).

Table 18: TEAEs during the double-blind study period reported for ≥2% in any treatment group (99)

Treatment-emergent adverse event	Daridorexant 50 mg N = 308 n (%)	Placebo N = 309 n (%)
Subjects with at least one event	116 (37.7)	105 (34.0)
Nasopharyngitis	20 (6.5)	20 (6.5)
Headache	19 (6.2)	12 (3.9)
Accidental overdose	8 (2.6)	5 (1.6)
Fatigue	7 (2.3)	2 (0.6)
Dizziness	7 (2.3)	2 (0.6)
Nausea	7 (2.3)	3 (1.0)

*Total number of subjects per treatment group with at least one event. Table is truncated to show only those AEs reported for at least 2% in any treatment group.

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row.
Includes TEAEs occurring (i.e., that started or worsened) during the double-blind study period.
AE = adverse event; MedDRA = Medical Dictionary for Regulatory Activities; TEAE = treatment emergent adverse event.

B.2.5.2 Subgroup analyses of TEAEs

Subgroup analysis was performed to evaluate treatment safety across the following demographic subgroups (99):

- Age: < 65, ≥ 65 years and < 75, ≥ 75 years
- Sex: Male, female
- BMI: 25, 25–30, > 30 kg/m^2

There were no clinically relevant differences between the two treatment groups in the overall incidence of TEAEs by age, sex or BMI (Appendix F).

B.2.5.3 Treatment-emergent SAEs

The incidence of treatment-emergent SAEs was low and were reported in 10 subjects: 3 (1.0%) and 7 (2.3%) subjects in the daridorexant 50 mg and placebo group, respectively (Table 19) (99).

Table 19: Treatment-emergent SAEs reported at least once in either treatment group (99)

Treatment-emergent SAE	Daridorexant 50 mg N = 308 n (%)	Placebo N = 309 n (%)
Subjects with at least one event	3 (1.0)	7 (2.3)
Syncope	1 (0.3)[a]	2 (0.6)
Adenocarcinoma of colon	1 (0.3)	0
Haemoglobin decreased	1 (0.3)[a]	0
Post procedural haemorrhage	1 (0.3)[a]	0
Renal colic	1 (0.3)*	0
Depression	0	2 (0.6)[b,*]
Anal abscess	0	1 (0.3)
Ankle fracture	0	1 (0.3)
Herpes zoster	0	1 (0.3)
Panic attack	0	1 (0.3)[b]

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row.
Preferred terms are based on MedDRA dictionary version 22.1.
Includes all SAEs occurring from start of double-blind study treatment up to 30 days after the end of double-blind study treatment or enrolment in the ID-078A303 extension study.
[a]Syncope, haemoglobin decreased, and post procedural haemorrhage were all reported for one subject.
[b] Depression and panic attack were both reported in the same subject.
*Renal colic and 1 of the 2 SAEs of depression occurred during the safety follow-up period.
SAE=Serious adverse event; MedDRA = Medical Dictionary for Regulatory Activities.

AEs leading to premature discontinuation of double-blind study treatment

AEs leading to premature study treatment discontinuation were reported for 3 (1.0%) and 10 subjects (3.2%) in the daridorexant 50 mg, and placebo groups, respectively (99).

B.2.5.4 AESIs

Incidence of treatment-emergent adverse events of special interest (AESIs) was low, with AESIs reported for 3 subjects (2 in daridorexant 50 mg], 1 in placebo). All AESIs were adjudicated as potentially related to study treatment by the ISB (Table 20) (99).

1. 'Narcolepsy-like symptoms related to excessive daytime sleepiness' were equally distributed across both treatment groups (1 subject each in the daridorexant 50 mg and placebo groups).

2. 'Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations and sleep paralyses' were reported for 1 subject in the daridorexant 50 mg group and none in the placebo group.

All adjudicated AESIs were non-serious, and the majority were of mild intensity, except for 2 events of moderate somnolence and 1 event of severe sleep paralysis. None of the events required treatment, and study treatment continued in all but 1 subject (99).

Table 20: Treatment-emergent AESIs after ISB adjudication (99)

Adverse event of special interest	Daridorexant 50 mg N = 308 n (%)	Placebo N = 309 n (%)
Subjects with at least one event	2 (0.6)	1 (0.3)
Narcolepsy-like symptoms related to excessive daytime sleepiness	1 (0.3)	1 (0.3)
Somnolence	1 (0.3)	1 (0.3)
Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations/sleep paralysis	1 (0.3)	0
Sleep paralysis	1 (0.3)	0

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA dictionary version 22.1
Includes all AESIs, as confirmed by ISB adjudication, occurring from start of double-blind study treatment up to 30 days after the end of double-blind study treatment or enrolment in the ID-078A303 extension study.
AESI = adverse event of special interest; ISB = Independent Safety Board; MedDRA = Medical Dictionary of Regulatory Activities.

B.2.5.5 Other safety assessments

Withdrawal symptoms

Withdrawal symptoms were assessed based on Benzodiazepine Withdrawal Symptom Questionnaire (BWSQ) total score, AEs, and electrocardiogram (ECG) findings occurring during the placebo run-out period. Observations were comparable between the daridorexant 50 mg group and the placebo group and there was no trend suggestive of withdrawal-related symptoms upon discontinuation of daridorexant. No statistical comparisons were performed for all safety assessments (99).

Mean BWSQ scores were comparable between the daridorexant and placebo groups. Minor numerical change from the last assessment (during double-blind treatment) to placebo run-out period at visit 9 (mean reduction -0.3 [SD 1.7] for daridorexant 50 mg and -0.5 [1.8] for placebo group) and visit 10 (mean reduction -0.6 [2.3] for daridorexant 50 mg and -0.7 [2.3]) were observed (Table 21). During the placebo run-in period, 2.2% of subjects in the daridorexant 50 mg group and 1.1% of subjects in placebo group reported BWSQ total scores of >20. No subjects had a BWSQ total score of >20 at the end of the placebo run-out period (visit 10). The proportion of subjects with at least one symptom scored as severe on the BWSQ was highest during the placebo run-in period (11.9% in daridorexant 50 mg and 13.7% in placebo group), becoming progressively lower at month 1 and 3, and lowest at the end of placebo run-out period (4.3% in daridorexant 50 mg and 2.2% in placebo group) (99).

Table 21: Observed value and change in BWSQ total score from last available assessment of double-blind study treatment to each scheduled placebo run-out visit (99)

Time point Statistic	Daridorexant 50 mg N = 286	Placebo N = 280
Last assessment on double-blind treatment		
n	286	280
Mean (SD)	2.0 (3.1)	1.9 (3.3)
Run-out - Visit 9		
n	275	266
Mean (SD)	1.7 (2.7)	1.5 (2.7)
Change from last assessment on double-blind treatment to Run-out - Visit 9		
n	275	266
Mean (SD)	0.3 (1.7)	-0.5 (1.8)
Run-out - Visit 10		
n	282	273

Mean (SD)	1.4 (2.2)	1.2 (2.1)
Change from last assessment on double-blind treatment to Run-out - Visit 10		
n	282	273
Mean (SD)	-0.6 (2.3)	-0.7 (2.3)

BWSQ=Benzodiazepine Withdrawal Symptom Questionnaire; SD=standard deviation.

TEAEs during the placebo run-out period were reported for 8.0% and 6.1% of subjects in the daridorexant 50 mg and placebo groups, respectively. The most commonly reported TEAEs were nasopharyngitis and headache (99). No AEs suggestive of withdrawal symptoms were reported in both treatment groups (Table 22).

Table 22: Treatment-emergent AEs during placebo run-out reported in ≥2 subjects in either treatment group* (99)

Treatment-emergent adverse event	Daridorexant 50 mg N = 286 n (%)	Placebo N = 280 n (%)
Subjects with at least one event**	23 (8.0)	17 (6.1)
Nasopharyngitis	6 (2.1)	6 (2.1)
Headache	2 (0.7)	1 (0.4)
Accidental overdose	0	1 (0.4)
Influenza	0	2 (0.7)

* Includes all AEs occurring from the start of the run-out period until the end of the run-out period. The start of the run-out period is 1 day after the start of run-out placebo treatment (if the treatment was taken before midnight) or the day of the start of run-out placebo treatment (if the treatment was taken after midnight); the end of the run-out period is the latter of 7 days after the start of run-out period or the Visit 10 date.
**Total number of subjects per treatment group with at least one event. Table is truncated to show only those AEs reported for at least 2 subjects (0.7%) in any treatment group.
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row.
AE=adverse event

Incidence of marked ECG abnormalities during the placebo run-out period were [redacted] ([redacted]) (Table 27) (99).

Table 23: Marked ECG abnormalities during placebo run-out (99)

ECG parameter Statistic	Daridorexant 50 mg N = 286 n / Nn (%)	Placebo N = 280 N / Nn (%)
ECG Mean Heart Rate (beats/min)		
<50	[redacted]	[redacted]
>45	[redacted]	[redacted]
>10 and ≤20 decrease from baseline	[redacted]	[redacted]
>20 decrease from baseline	[redacted]	[redacted]
PR Interval, Single Beat (msec)		
>200	[redacted]	[redacted]
QRS Duration, Single Beat (msec)		

>110	[redacted]	[redacted]
QTcB Interval, Single Beat (msec)		
>450 and ≤ 480	[redacted]	[redacted]
>480 and ≤ 500	[redacted]	[redacted]
> 30 and ≤ 60 increase from baseline	[redacted]	[redacted]
QTcF Interval, Single Beat (msec)		
>450 and ≤ 480	[redacted]	[redacted]
> 30 and ≤ 60 increase from baseline	[redacted]	[redacted]

Nn is the number of subjects at risk: those having at least one post-baseline value per ECG parameter for criterion based on post-baseline values only, or those having a baseline value and at least one post-baseline value per ECG parameter for criterion based on change from baseline.
ECG = electrocardiogram; QTc = QT interval corrected for heart rate; QTcB = QT interval corrected for heart rate according to Bazett's formula; QTcF = QT interval corrected for heart rate according to Fridericia's formula.

Drug abuse potential

AEs related to drug abuse, dependence and withdrawal were also studied in subjects of study 301. TEAEs related to abuse were reported for 14 (4.5%) and 11 subjects (3.6%) in the daridorexant 50 mg group and placebo group, respectively. Accidental overdose was reported for 8 subjects (2.6%) in the daridorexant 50mg group, and 5 subjects (1.6%) in the placebo group (99). There were no AEs reported for intentional overdose, while reports of overdose (unspecified) or accidental overdose were asymptomatic with no evidence of withdrawal symptoms (Appendix F).

Sheehan disability scale© (SDS©)

Overall, the SDS© scores showed no signs of impaired daytime functioning on any of the assessed sub-scores related to daridorexant, and were comparable across both treatment groups. In both treatment groups, mean total SDS© scores decreased from baseline to month 1 [redacted], month 3 [redacted] and placebo run-out period [redacted] (Table 24) (99).

Table 24: Sheehan Disability Scale© total score – Observed value and change from baseline at Month 1, Month 3 and placebo run-out (99)

Time point Statistic	Daridorexant 50 mg N = 286	Placebo N = 280
Baseline		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Month 1		

Time point Statistic	Daridorexant 50 mg N = 286	Placebo N = 280
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Change from baseline to Month 1		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Month 3		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Change from baseline to Month 3		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Run-out - Visit 9		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Change from baseline to Run-out – Visit 9		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]

Values for Month 1 / Month 3 / run-out were calculated only for subjects who had a baseline value.
A decrease in Sheehan Disability Scale© score indicates an improvement.
SD=standard deviation.

The number of days reported as lost in the week prior to the assessment [redacted] treatment groups at baseline ([redacted] in the daridorexant 50 mg, and placebo groups, respectively) (99). [redacted] across all timepoints (month 1, month 3, and placebo run-out) [redacted] daridorexant 50 mg over placebo (Table 25). The [redacted] in days lost was observed at month 3 [redacted] (99).

Table 25: Sheehan Disability Scale© – Observed value and change from baseline in number of days lost in a week at month 1, month 3 and placebo run-out (99)

Time point Statistic	Daridorexant 50 mg (N = 308)			Placebo (N = 309)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Month 1						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Month 3						

Time point Statistic	Daridorexant 50 mg (N = 308)			Placebo (N = 309)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Run-out – Visit 9						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SD=Standard Deviation.

The number of days reported as underproductive in the week prior to the assessment was [redacted] groups at baseline ([redacted] in the daridorexant 50 mg, and placebo groups, respectively). Reduction in unproductive days [redacted] of daridorexant 50 mg was observed at month 1, month 3, and placebo run-out (Table 26), with the [redacted] difference observed at placebo run-out [redacted] (99).

Table 26: Sheehan Disability Scale© – Observed value and change from baseline in number of unproductive days in a week at Month 1, Month 3 and placebo run-out (99)

Time point Statistic	Daridorexant 50 mg (N = 308)			Placebo (N = 309)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Month 1						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Month 3						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Run-out – Visit 9						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SD=Standard Deviation.

Additional safety assessments of vital signs (mean of the 2 PSG nights in systolic and diastolic BP and pulse rate), change in body weight, marked laboratory abnormalities, occurrence of suicidal ideation and/or behaviour, next-day residual effects, and rebound insomnia are reported in Appendix F.

B.2.6 Study 303 — summary of trial methodology

B.2.6.1 Study sites

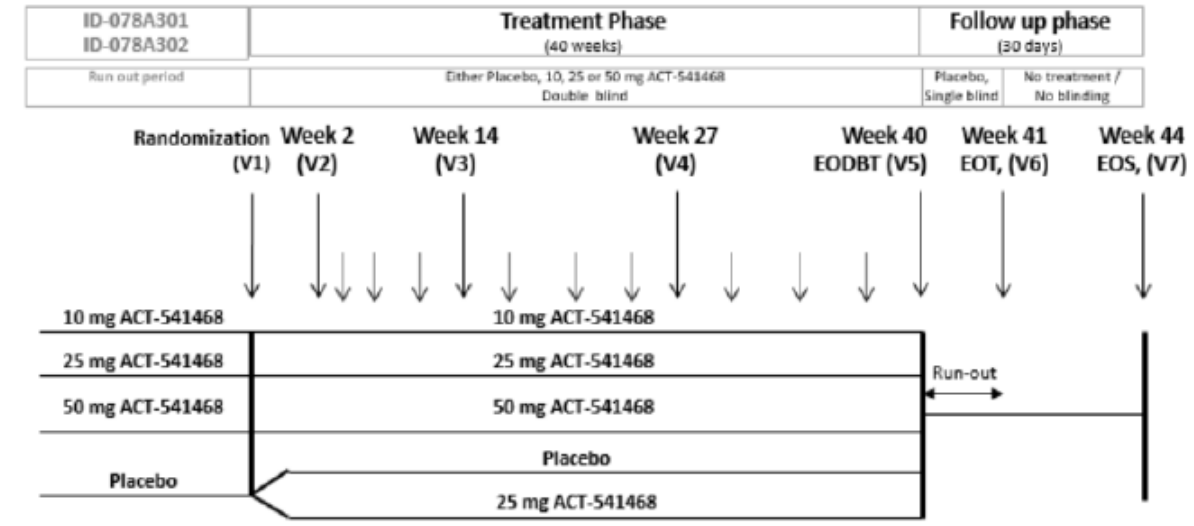
Study 303 was conducted in 94 sites across 14 countries (Belgium, Bulgaria, Canada, Denmark, Finland, France, Germany, Hungary, South Korea, Poland, Spain, Sweden, Switzerland, and the US) (97).

B.2.6.2 Study design

Study 303 was a multi-centre, double-blind, parallel-group, randomised, placebo- controlled, extension of confirmatory studies 301 and 302. Subjects who completed double-blind study treatment and the placebo run-out period of the confirmatory studies 301 and 302, and who were willing to participate, were eligible to enrol into study 303. Subjects assigned to a daridorexant group in study 301 or 302 were assigned to the same daridorexant dose (i.e., 10 mg, 25 mg or 50 mg). Subjects assigned to the placebo group in studies 301 or 302 were re-randomized to receive either placebo or daridorexant 25 mg in a 1:1 ratio in study 303 (97).

The overall study design is presented in Figure 11.

Figure 11: Design of study 303 (97)



·····> Monthly telephone call.
V2 and V7 were telephone calls; all other visits were at the site.
EODBT = End of Double-Blind Treatment; EOS = End-of-Study; EOT = End-of-Treatment.
* Subjects randomly assigned to the ACT-541468 arms in the confirmatory studies receive the same dose in the extension study.
** Subjects randomly assigned to the placebo arm in the confirmatory studies were randomized to receive either placebo or 25 mg ACT-541468 in a 1:1 ratio in the extension study.

The study comprised of a treatment phase and a safety follow-up phase (97):

- **Treatment phase:** from signing informed consent (Visit 1) until 40 weeks (Visit 5). Visit 1 was performed on the same day as EOT of the 301 or 302 study, after the placebo run-out assessments had been completed, or as an independent visit within a maximum of seven days after EOT. The treatment phase started with double-blind treatment allocation.

- A safety telephone call was performed any day from Day 7 to Day 14 (Visit 2) to collect information about AEs and concomitant medications. Safety parameters of each subject were assessed at Week 14 (Visit 3), Week 27 (Visit 4) and Week 40 (Visit 5). The SDQ and IDSIQ were completed at home daily during the last week of each consecutive four-week period during the double-blind treatment phase. Reminder telephone calls were scheduled within one week prior to questionnaire completion to ensure subject's compliance. EODBT was reached at Week 40 (Visit 5). Subjects who prematurely discontinued study treatment but remained in the study were encouraged to continue with all planned study procedures until EOS, excluding the placebo run-out period.

- **Safety follow-up phase:** from EODBT until EOS (30-day follow-up telephone call), comprising of a single-blind placebo run-out period of 7 days and a safety follow- up period:

    - **Placebo run-out period:** started in the evening of the Week 40 (Visit 5) and ended with EOT at Week 41 (Visit 6). Single-blind placebo treatment was taken once daily at bedtime. The SDQ and IDSIQ were completed daily at home.

    - **Safety follow-up period:** started after the placebo run-out period and ended 30 days after the last dose of double-blind study treatment with the 30-day follow-up telephone call (Week 44, Visit 7), which collected information on AEs, SAEs and concomitant medications.

B.2.6.3 Study eligibility criteria

Table 27 presents the key inclusion and exclusion criteria of extension study 303 (97).

Table 27: Inclusion and exclusion criteria of study 303 (97)

Inclusion criteria	• Signed informed consent prior to any study-mandated procedure (Visit 1). • Completion of the double-blind study treatment and placebo run-out period of 301 or 302 (Visit 1). • For woman of childbearing potential, the following was required: o Negative urine pregnancy test (EOT of 301 or 302 studies). o Agreement to use the contraception scheme as required by the protocol from Visit 1 up to at least 30 days after EODBT.
Exclusion criteria	• Subjects self-reporting daytime napping ≥ 1 hour per day and ≥ 3 days per week. • Subjects with BMI < 18.5 or > 40.0 kg/m^2. • Subjects who were pregnant, breastfeeding, or planning to become pregnant. • Subjects with any lifetime history of suicide attempt, sleep-related breathing disorders, periodic limb movement disorder, restless legs syndrome, circadian rhythm disorder, REM behaviour disorder, narcolepsy, or apnoea/hypopnea. • Subjects with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments. • Subjects aged ≥ 50 years with a Mini Mental State Examination© score < 25. • Subjects treated with central nervous system-active drugs; cognitive behavioural therapy was allowed if started at least 1 month prior to the run-in PSG nights and intended to be continued throughout the study. • Subjects not able or willing to stop treatment with moderate or strong CYP3A4 inhibitors or inducers within at least 1 week prior to the start of the placebo run-in period.

CYP3A4=cytochrome P450 3A4; EOT=End of treatment; EODBT=End-of-double-blind treatment; REM=rapid eye movement.

B.2.6.4 Study treatment, prior and concomitant medications

Study treatment comprised double-blind treatment (daridorexant and placebo matching daridorexant) administered from study treatment allocation to EODBT and single-blind treatment (placebo matching daridorexant) administered during the placebo run-out period (Table 28) (97).

Table 28: Trial drugs in study 303 (97)

Drug	Dose	Frequency of administration	Route of administration	Duration
Daridorexant, film coated tablet	10 mg, 25 mg and 50 mg	One tablet taken orally once daily in the evening	Oral	280 ± 7 days
Placebo matching daridorexant, film coated tablet	-			Treatment period (280 ± 7 days), and Single-blind placebo run-out period (7 + 2 days)

Therapies considered necessary for the subject's well-being and not categorized as prohibited concomitant medications could be used in the study, including COVID-19 vaccines (97).

The following concomitant therapies were forbidden during the study (97):

- Treatment with another investigational drug until EOS
- Study-prohibited CNS-active medications from at least 1 week prior to Visit 1 and until 24 hours after EOT
- Treatment with moderate or strong CYP3A4 inhibitors or moderate to strong CYP3A4 inducers from at least 1 week prior to Visit 1 until 24 hours after EOT.

B.2.6.5 Pre-specified study endpoints

The pre-specified endpoints relevant for the decision problem are summarised in Table 29.

Table 29: Primary and exploratory endpoints of study 303 relevant to NICE decision problem (97)

Primary endpoint	Definition	NICE scope/ economic model?
Safety	*The statistical methods and results of the following safety endpoints are reported in B.2.6.6 and B.2.8, respectively.* • TEAEs up to 30 days after double-blind study treatment discontinuation. • SAEs up to 30 days after study double-blind treatment discontinuation. • AEs leading to premature discontinuation of the double-blind study treatment. • AESIs after adjudication by ISB: o Narcolepsy-like symptoms (i.e., EDS, cataplexy and complex sleep behaviour events including hallucinations/sleep paralysis) o Suicide/self-injury. • Withdrawal effects (physical dependence) upon treatment discontinuation, assessed based on the changes from last assessment on double-blind treatment (Visit 5, Week 40) to end of the placebo run-out period (Visit 6, Week 41) in the BWSQ total score, the occurrence of relevant AEs and marked ECG abnormalities.	Per NICE scope, not included in the economic model

	• Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), and Visit 5 (Week 40) in SDS©.	
	The results of the following safety endpoints are reported in Appendix M. • Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), Visit 5 (Week 40), and Visit 6 (Week 41, run-out period) in ESS© total score • Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), and Visit 5 (Week 40) in vital signs (systolic and diastolic BP, and pulse rate). • Change from baseline[a] to Visit 5 (Week 40) in body weight. • Marked ECG abnormalities on double-blind study treatment. • Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), Visit 5 (Week 40), and Visit 6 (Week 41, run-out period) in ECG parameters. • Marked laboratory abnormalities on double-blind study treatment. • Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), Visit 5 (Week 40), and Visit 6 (Week 41, run-out period) in laboratory parameters. • Occurrence of suicidal ideation and/or behaviour on double-blind study treatment and during the placebo run-out period based on C-SSRS© • Rebound insomnia, assessed based on change from baseline[b] to the placebo run-out period[c] in sTST • Next-morning residual effect, assessed based on change from baseline[a] over time[d] in morning sleepiness score on the SDQ VAS (mm)	Per NICE scope, not included in the economic model
Exploratory endpoints		**NICE scope/ economic model?**
Other exploratory endpoints	*The statistical methods and results of ISI© are reported in B.2.6.6 and B.2.7.1, respectively.* • Change from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), and Visit 5 (Week 40) in ISI© scores • Number (%) of subjects with ≥6-point decrease in ISI© total score (100) from baseline[a] to Visit 3 (Week 14), Visit 4 (Week 27), and Visit 5 (Week 40)	Per NICE scope, included in the economic model
	The statistical methods and results of the following exploratory endpoints are reported in B.2.6.6 and B.2.7.1, respectively. • Change from baseline[b] over time[d] in sTST. sTST is the total sleep time as reported in answer to item 9 of the SDQ ('In total, how long did you sleep last night?'). • Change from baseline[b] over time[d] in sLSO. sLSO is the self-reported time to fall asleep as reported in answer to item 5 of the SDQ ('How long did it take you to fall asleep?').	Per NICE scope, not included in the economic model

	• Change from baseline[b] over time[d] in subjective sleep maintenance (sWASO). sWASO is the self-reported time spent awake after sleep onset as reported in answer to item 7 of the SDQ ('In total, how long did these awakenings last?'). • Change from baseline[b] over time[d] in IDSIQ scores (i.e., total score; alert/cognition, mood and sleepiness domain scores).	
	The results of the following exploratory endpoints are reported in Appendix M. • Change from baseline[a] over time[c] in scores on the SDQ VAS (mm). VAS scores are the subjects' ratings of 'the quality of your sleep last night', 'the depth of your sleep last night', daytime alertness (from 'your daytime alertness today') and daily ability to function (from 'your daily ability to function today') questions. • Change from baseline[a] over time[c] in mean number of self-reported awakenings. The number of self-reported awakenings is the number reported in answer to item 6 of the SDQ ('How many times did you wake up, not counting your final awakening?'). • Change from baseline[a] over time[c] in PGA-S scores (daytime symptoms) • Change from baseline[a] over time[c] in PGI-C scores (daytime symptoms)	Per NICE scope, not included in the economic model

[a]Baseline refers to: 'confirmatory study baseline' for daridorexant 50 mg and for placebo; 'extension study baseline' for the ex-placebo/daridorexant 25 mg group.
[b]Baseline refers to: 'confirmatory study baseline' for daridorexant 50 mg; 'extension study baseline' for the ex-placebo/daridorexant 25 mg group; for the placebo group, change from baseline was analysed in 2 ways: change from the 'confirmatory study baseline' (as planned) and change from the 'extension study baseline' (added after unblinding for interim analysis).
[c]Run-out period: the mean value of the SDQ entries in the 7 consecutive days immediately following the evening of Visit 5.
[d]Over time: the mean value of the SDQ entries for each week in which this questionnaire was completed (except for the run-out week).
AEs=Adverse events; AESI= Adverse event of special interest; BWSQ= Benzodiazepine Withdrawal Symptom Questionnaire; C-SSRS©= Columbia Suicide Severity Rating Scale©; ECG= electrocardiogram; EDS=excessive daytime sleepiness; ESS©= Epworth Sleepiness Scale©; IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISB=independent safety board; ISI©=Insomnia severity index©; LPS=latency to persistent sleep; PGA-S=Patient Global Assessment of Disease Severity; PGI-C=Patient Global Impression of Change; PGI-S=Patient Global Impression of Severity PICO=population, intervention, comparator and outcome; PSG=polysomnography; REM=rapid eye movement; S1, S2, S3= sleep stage 1, 2 and 3; SAE=Serious AEs; SDS= Sheehan Disability Scale©; SDQ=Sleep diary questionnaire; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake after sleep onset; SWS= slow-wave sleep; TEAEs= Treatment-emergent AEs; TST= total sleep time; VAS=visual analogue scale; WASO=wake after sleep onset.

B.2.6.6 Statistical methods and analysis sets

Table 30 details the statistical methods and analysis sets used in extension study 303 (97).

Table 30: Summary of statistical methods and analysis sets of study 303 (97)

Study name (number)	Study 303 (NCT03679884)
Analysis sets	• **Enrolled set:** The Enrolled set included all subjects who completed study 301 or 302 and who consented to enter study 303. • **Full analysis set:** The Full analysis set comprised all subjects assigned (i.e., randomized) to a study treatment.

Study name (number)	Study 303 (NCT03679884)
	• **Safety set:** The Safety set comprised all subjects who received at least one dose of double-blind study treatment. • **Treatment withdrawal set:** The Treatment withdrawal set comprised all subjects in the Safety set who received at least one dose of single-blind placebo treatment in the placebo run-out period.
Statistical analysis of safety endpoints	All safety endpoints were summarised descriptively.
Statistical analysis for exploratory efficacy endpoints	Analysis of exploratory efficacy endpoints was performed using the Full analysis set. • Linear mixed effects model was used for the analysis of change from confirmatory baseline in sTST, sWASO, sLSO and IDSIQ total score, sleepiness domain, alert/cognition domain, and mood domain scores, separately. • The analysis model adjusted for the confirmatory baseline value of the relevant response variable (either sWASO, sLSO, sTST, or IDSIQ total score, sleepiness domain, alert/cognition domain, or mood domain scores), age group as per assigned strata (< 65; ≥ 65 years), treatment (daridorexant 50 mg; placebo), visit (at Month 6 [Week 12 of extension study]; Month 9 [Week 24]; Month 12 [Week 36]), and the interaction of treatment by visit, and baseline by visit. • Appropriate contrasts were used to test the difference in LSM change from confirmatory baseline between daridorexant 50 mg and placebo at Month 6 [Week 12]; Month 9 [Week 24]; and Month 12 [Week 36]. Observed values and change from baseline over time in ISI© were summarized descriptively.
Sample size & power calculation	As study 303 was an extension of studies 301 and 302, no formal sample size calculation was undertaken. It was expected that approximately 1,260 subjects (i.e., ~ 70% of the total subjects in studies 301 and 302) would enter the extension study, assuming all sites participated in this study.
Data management, patient withdrawals	**Handling of missing data:** For sTST, sWASO, sLSO, each IDSIQ domain and total scores, VAS scores and number of self-reported awakenings, at least 2 days of data during each week were required to calculate a weekly mean. Otherwise, the mean value was considered missing for that week. The approach implies implicit imputation: the missing data points were given the same value as the mean of the non-missing data points of that same time point or week.

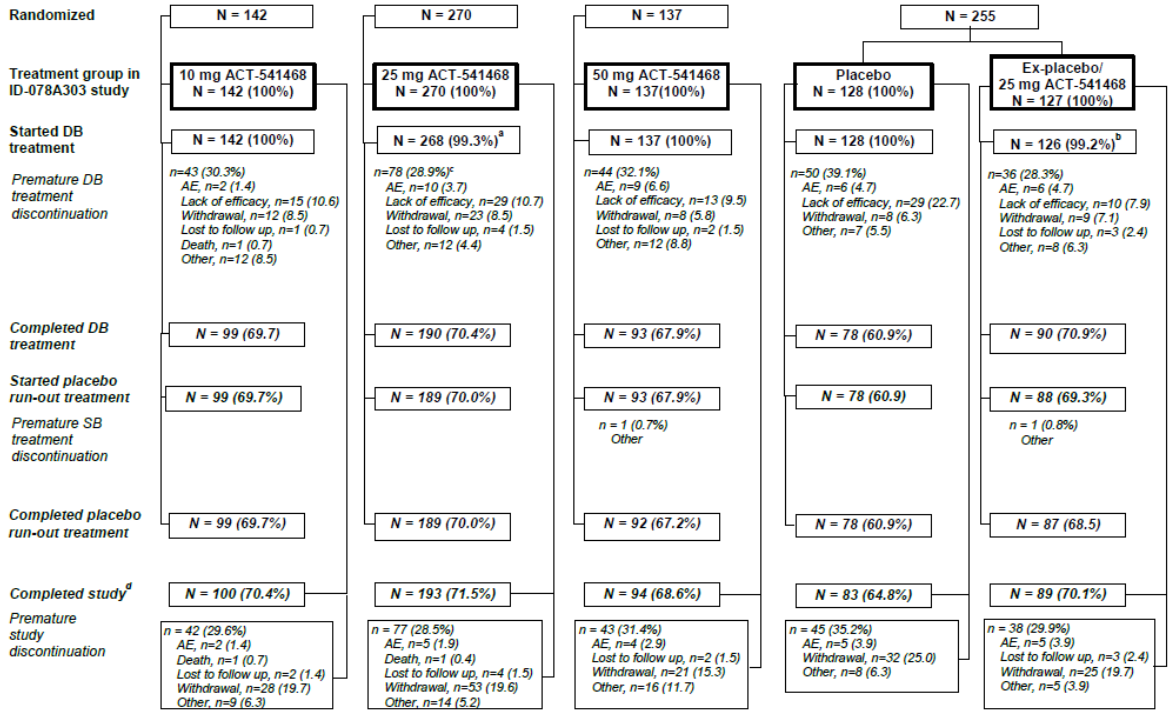
IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; ISI©= Insomnia severity index©;; LS=least squares; sLSO=subjective latency to sleep onset; sTST=subjective total sleep time; sWASO=subjective wake after sleep onset;; VAS=visual analogue scale.

B.2.6.7 Participant flow

In total, 804 subjects entered the extension study 303; 801 subjects started treatment; three subjects discontinued from the study prior to starting the treatment (Figure 12). Subjects who received any dose of daridorexant in studies 301 and 302 were re- randomised to receive 25 mg daridorexant or placebo in study 303. A total of 137 subjects received daridorexant 50 mg, while 128 subjects received placebo in study

303 (97). Of the 804 randomized subjects, 251 (31.2%) prematurely discontinued double-blind treatment. The most frequent reason for this premature discontinuation was lack of efficacy (11.9% overall), which was more frequent in the placebo group (22.7%) compared to daridorexant treatment groups (97).

Figure 12: Disposition of subjects in study 303 (97)



Percentages are calculated out of subjects who were randomized.
[a]2 Subjects (Subject 1908175 and Subject 3540013) withdrew from the study prior to receiving double-blind treatment
[b]1 Subject (Subject 1902066) discontinued from the study prior to receiving double-blind treatment due to a positive urine drug test.
[c]For 1 subject (Subject 1912065), the reason for discontinuation of double-blind treatment was recorded as "other: medical reasons following SAE" rather than "AE"
[d]Subjects completed the 30-day follow-up telephone call.
AE=adverse event; DB=double-blind; SAE=serious adverse event; SB=single-blind.

B.2.6.8 Baseline characteristics and demographics

Demographic characteristics were overall balanced across treatment groups in the Full analysis set of extension study 303, similar to the confirmatory studies of 301 or 302 (Table 31).

Table 31: Demographic characteristics of subjects enrolled in study 303 (97)

Variable Statistic	Daridorexant 50 mg N=137	Placebo N=128
Age at screening (years)		
Mean (SD)	56.9 (13.6)	59.2 (12.6)
Median (Min, Max)	59 (22, 81)	61 (30, 85)
Sex [n(%)]		
Male	39 (28.5)	36 (28.1)

Variable Statistic	Daridorexant 50 mg N=137	Placebo N=128
Female	98 (71.5)	92 (71.9)
Race [n(%)]		
Black or African American	15 (10.9)	8 (6.3)
American Indian or Alaska Native	1 (0.7)	0
Native Hawaiian or other Pacific Islander	0	1 (0.8)
Asian	0	2 (1.6)
White	121 (88.3)	115 (89.8)
Other	0	2 (1.6)
Ethnicity [n(%)]		
Hispanic or Latino	19 (13.9)	10 (7.8)
Not Hispanic or Latino	118 (86.1)	118 (92.2)
BMI (kg/m^2) at screening		
Mean (SD)	25.890 (4.238)	25.904 (4.039)
Region [n(%)]		
US	36 (26.3)	46 (35.9)
Other (non-US)	101 (73.7)	82 (64.1)

*All demographic data reported in this table are from the respective confirmatory study 301
BMI=Body mass index; SD=standard deviation; US=United States.

Baseline values of efficacy variables were well balanced across treatment groups for subjects who remained on the treatment during the extension study they were assigned to in the confirmatory studies, using the confirmatory study baseline (97).

Table 32: Baseline sTST, sLSO, sWASO, IDSIQ total and domain scores (97)

Time point statistic	Daridorexant 50 mg	Placebo
	N = 137	N = 128
sTST (min)		
n	137	128
Mean (SD)	303.792 (65.084)	305.071 (56.506)
IDSIQ sleepiness domain score		
n	137	128
Mean (SD)	22.374 (6.562)	21.792 (6.564)
IDSIQ total score		
n	137	128
Mean (SD)	74.864 (23.519)	70.297 (22.125)
IDSIQ alert/cognition domain score		
n	137	128
Mean (SD)	32.389 (9.999)	30.826 (9.138)
IDSIQ mood domain score		
n	137	128
Mean (SD)	20.101 (8.014)	17.679 (8.005)
sLSO (min)		
n	137	128

Time point statistic	Daridorexant 50 mg	Placebo
	N = 137	N = 128
Mean (SD)	63.409 (40.300)	64.821 (39.952)
sWASO (min)		
n	137	128
Mean (SD)	80.114 (57.327)	82.675 (52.388)

IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; SD = standard deviation; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; sWASO = subjective wake after sleep onset.

B.2.7 Study 303 — clinical effectiveness results

B.2.7.1 Exploratory efficacy endpoints

Since the primary aim of study 303 was to assess the long-term safety and tolerability of daridorexant, all efficacy endpoints were exploratory. Subjects treated with daridorexant 50 mg [redacted] in sTST compared to placebo at month 6 [redacted], but not at months 9 [redacted] and 12 [redacted] (Table 33) (97).

Table 33: Exploratory endpoints – sTST (min) (97)

Visit	n	LSM [95% CL]	Difference to placebo	
Treatment group			LSM [95% CL]	p-value (two-sided)
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	105	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	98	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	97	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	80	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	87	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	70	[redacted]	[redacted]	[redacted]

CL=confidence limit; LSM=least squares mean; sTST=subjective total sleep time.

Month 6 time point includes the duration of the confirmatory study and corresponds to Week 12 of the extension study, same for Month 9 (Week 24) and Month 12 (Week 36).
Mixed effects model for Repeated Measures: Change from baseline in sTST = baseline sTST + stratified age group (< 65; >= 65 years) + treatment + visit + treatment * visit + baseline * visit.
n is the number of subjects with non-missing values.

Subjects in the daridorexant 50 mg group reported [redacted] from baseline in IDSIQ sleepiness domain score compared to placebo at month 6, month 9 and month 12

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx. Likewise, IDSIQ total, alert/cognition domain and mood domain scores xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx in the daridorexant 50 mg group compared to placebo at month 6, month 9 and month (Table 34) (97).

Table 34: Exploratory endpoints – IDSIQ sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score (97)

Visit / Treatment group	n	LSM [95% CL]	Difference to placebo: LSM [95% CL]	Difference to placebo: p-value (two-sided)
Between treatment analysis for change from baseline in IDSIQ Sleepiness domain score to Month 6, Month 9 and Month 12				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ total score to Month 6, Month 9 and Month 12				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ alert/cognition domain score to Month 6, Month 9 and Month 12				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				

Visit Treatment group	n	LSM [95% CL]	Difference to placebo	
			LSM [95% CL]	p-value (two-sided)
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ mood domain score to Month 6, Month 9 and Month 12				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]

Higher IDSIQ score represents greater burden of illness.
IDSIQ=Insomnia Daytime Symptoms and Impacts Questionnaire; LSM=least squares mean; SD=standard deviation.
Month 6 timepoint includes the duration of the confirmatory study and corresponds to Week 12 of the extension study, same for Month 9 (Week 24) and Month 12 (Week 36).
Mixed effects model for Repeated Measures: Change from baseline in IDSIQ Sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score = baseline IDSIQ Sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score + stratified age group (< 65; >= 65 years) + treatment + visit + treatment * visit + baseline * visit.

Compared to placebo, xxxxxxxxxxxxx in daridorexant 50 mg for sLSO from confirmatory study baseline was observed at month 6 [redacted], at month 9 [redacted] and at month 12 [redacted] (97). Similarly, no reductions in sWASO from confirmatory study baseline in the daridorexant 50 mg group compared to placebo were observed at month 6 [redacted], month 9 [redacted] and month 12 [redacted] (Table 35).

Table 35: Exploratory endpoints – sLSO (min) and sWASO (min) (97)

Visit / Treatment group	n	LS Mean [95% CL]	Difference to placebo	
			LS Mean [95%CL]	p-value (two-sided)
Between treatment analysis for change from baseline in sLSO (min) to Month 6, Month 9 and Month 12, Full analysis set				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in sWASO (min) to Month 6, Month 9 and Month 12, Full analysis set				
Change from baseline to Month 6				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 9				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline to Month 12				
Daridorexant 50 mg (N = 137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N = 128)	[redacted]	[redacted]	[redacted]	[redacted]

LSM=Least squares mean; SD=standard deviation; sLSO=subjective latency to sleep onset; sWASO=subjective wake after sleep onset.

Numerically, mean ISI© scores [redacted] from [redacted] at confirmatory study baseline to [redacted] at week 27 [redacted] and [redacted] at week 40 [redacted] for daridorexant 50 mg group compared to placebo (97) (Table 36). Responder analysis revealed [redacted] reporting a decrease in ISI© score of ≥6 points from confirmatory study baseline in the daridorexant group compared to placebo across all timepoints (Table 37).

Table 36: Exploratory endpoint – ISI© score

Time point Statistic	Daridorexant 50 mg (N = 137)			Placebo (N = 128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Week 14						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 27						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 40						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ISI©=Insomnia severity index; SD=Standard Deviation.

Table 37: Exploratory endpoint – Subjects with ≥6 points decrease in total score from baseline to week 14, week 27 and week 40

Time point	Daridorexant 50 mg N = 137 n/ Nn (%)	Placebo N = 128 n/ Nn (%)
Week 14	[redacted]	[redacted]
Week 27	[redacted]	[redacted]
Week 40	[redacted]	[redacted]

Nn is the number of subjects with non-missing values at the given scheduled visit.

B.2.7.2 Subgroup analyses of exploratory efficacy endpoints

Subgroup analyses of sTST, sWASO, sLSO, IDSIQ domain and total scores were performed to investigate the consistency of the treatment effect across the following subgroups (97):

- Age at screening of confirmatory study: < 65, ≥ 65 years and < 75, ≥ 75 years.

Additionally, the following subgroup analyses were performed for sTST and IDSIQ domain and total scores:

- Sex: Male, female.
- Region: US, other (non-US).
- BMI at screening of confirmatory study: < 30, ≥ 30 kg/m^2.

- Race: White, Black or African American.

[redacted] with the subgroup analysis performed in confirmatory study 301, there were [redacted] in treatment effect across all subgroups as shown by the [redacted] (Appendix E). Overall, the [redacted] with that of the overall population in extension study 303 (97).

B.2.8 Study 303 — adverse reactions

B.2.8.1 TEAEs

During the double-blind study period, 38.0% and 33.6% of subjects reported TEAEs in the daridorexant 50 mg and placebo groups, respectively. Most of the events were of mild or moderate intensity (97). Nasopharyngitis [8.0% (daridorexant 50 mg); 4.7% (placebo)] was the most commonly reported TEAE (Table 38). Additional TEAEs with an incidence of ≥ 2% in both daridorexant 50mg and placebo groups were accidental overdose (2.9% vs 0%), somnolence (2.9% vs 0%), cough (2.2% 0%) and pneumonia (2.2% vs 0%) (Table 38) (97).

TEAEs that occurred during the double-blind study period considered [redacted] were reported [redacted] in the daridorexant 50 mg, and placebo groups, respectively (97). [redacted] was the [redacted] TEAE considered related to study treatment in [redacted] (97).

Table 38: TEAEs during the double-blind study period reported for ≥2% of subjects in either treatment group* (97)

Treatment-emergent adverse event	Daridorexant 50 mg N = 137 n (%)	Placebo N = 128 n (%)
Subjects with at least one event**	52 (38.0)	43 (33.6)
Nasopharyngitis	11 (8.0)	6 (4.7)
Accidental overdose	4 (2.9)	0
Somnolence	4 (2.9)	0
Fall	3 (2.2)	2 (1.6)
Headache	3 (2.2)	2 (1.6)
Cough	3 (2.2)	0
Pneumonia	3 (2.2)	0

*Includes only those TEAEs occurring (i.e., that started or worsened) during the double-blind study period.
**Total number of subjects per treatment group with at least one event. Table is truncated to show only those AE PTs reported for at least 2% in any treatment group.
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA version 22.1
AE = adverse event; PT=Preferred terms; TEAE = treatment-emergent adverse event.

B.2.8.2 Subgroup analyses of TEAEs

Subgroup analyses was performed to evaluate treatment safety across the following demographic subgroups (97):

- Age: < 65, ≥ 65 years and <75, ≥75 years
- Sex: Male, female
- BMI: < 25, 25–30, > 30 kg/m^2
- Race: White, Black or African American, Other

There [redacted] in the overall incidence of TEAEs by age, sex, BMI or race (Appendix F).

B.2.8.3 Treatment-emergent SAEs

The incidence of treatment-emergent SAEs was low (5.1% subjects in the daridorexant 50 mg group vs 1.6% subjects in the placebo group) (Table 39) (97).

Table 39: Treatment-emergent SAEs reported at least once in either treatment group (97)

Treatment-emergent SAE	Daridorexant 50 mg N = 137 n (%)	Placebo N = 128 n (%)
Subjects with at least one event	7 (5.1)	2 (1.6)
Diverticulitis	1 (0.7)	0
Confusional state	1 (0.7)	0
Bone disorder	1 (0.7)	0
Chronic lymphocytic leukaemia	1 (0.7)	0
Influenza like illness	1 (0.7)	0
Pneumonia	1 (0.7)	0
Thyroiditis subacute	1 (0.7)	0
Wrist fracture	1 (0.7)	0
Depression	0	1 (0.8)
Head injury	0	1 (0.8)

Treatment-emergent SAE	Daridorexant 50 mg N = 137 n (%)	Placebo N = 128 n (%)
Subdural haematoma	0	1 (0.8)
Suicidal ideation	0	1 (0.8)

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA version 22.1
Includes all AEs in the double-blind study period and up to 30 days after double-blind study treatment end date.
AE=Adverse event; SAE=Serious adverse event.

AEs leading to premature discontinuation of double-blind study treatment

AEs leading to premature study treatment discontinuation were reported for [redacted] in the daridorexant 50 mg and placebo groups, respectively (97).

B.2.8.4 AESIs

Incidence of treatment-emergent AESIs [redacted], with AESIs reported for [redacted]. All AESIs were [redacted] to study treatment by the ISB (Table 40) (97).

1. A [redacted] belonging to the category [redacted] related to complex sleep behaviour [redacted]
2. A [redacted] belonging to the [redacted]

There were [redacted] to the category [redacted] (97).

Table 40: Treatment-emergent AESIs after ISB adjudication (97)

Adverse event of special interest	Daridorexant 50 mg N = 137 n (%)	Placebo N = 128 n (%)
Subjects with at least one event	[redacted]	[redacted]
Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations/sleep paralysis	[redacted]	[redacted]
Abnormal dreams	[redacted]	[redacted]
Suicide/self-injury	[redacted]	[redacted]

Suicidal ideation	[redacted]	[redacted]

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row.
Includes all AEs in the double-blind study period and up to 30 days after double-blind study treatment end date.
AE=Adverse event; AESI = adverse event of special interest; ISB = Independent Safety Board.

B.2.8.5 Other safety assessments

Withdrawal symptoms

Overall, the analysis of BWSQ total score, AEs, and ECG findings during the placebo run-out period showed [redacted] withdrawal-related symptoms upon discontinuation of daridorexant. No statistical comparisons were performed for all safety assessments (97).

Mean BWSQ scores were [redacted] between the daridorexant and placebo groups. [redacted] from the last assessment (on double-blind treatment) to placebo run-out period ([redacted]) was observed (Table 41). [redacted] had a [redacted] at the end of the placebo run-out period. The [redacted] with at [redacted] scored as [redacted] on the BWSQ [redacted] during the last assessment of double-blind treatment ([redacted]), and almost [redacted] at the end of placebo run-out period [redacted] (97).

Table 41: Observed value and change in BWSQ total score from last available assessment of double-blind study treatment to end of placebo run-out (97)

Time point Statistic	Daridorexant 50 mg N = 93	Placebo N = 78
Last assessment on double-blind treatment value		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
End of run-out period		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Change from last value on double-blind treatment		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]

BWSQ=Benzodiazepine Withdrawal Symptom Questionnaire; SD=standard deviation.

AEs during the placebo run-out period were reported [redacted] in the daridorexant 50 mg and placebo groups, respectively. The most commonly reported TEAEs were [redacted]. [redacted] were reported in [redacted] (Table 42) (97).

Table 42: AEs during placebo run-out reported in ≥1 subject in either treatment group (97)

Preferred term	Daridorexant 50 mg N = 93 n (%)	Placebo N = 78 n (%)
Subjects with at least one event	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Concussion	[redacted]	[redacted]
Cystitis	[redacted]	[redacted]
Fall	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Hyperbilirubinemia	[redacted]	[redacted]
Migraine	[redacted]	[redacted]
Sciatica	[redacted]	[redacted]

Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row.
Adverse events which occur on or after the run-out period start until the run-out period end are displayed; Run-out period start is defined as one day after run-out single-blind placebo treatment start date (if the treatment is taken before midnight) or the day of run-out single-blind placebo treatment start date (if the treatment is taken after midnight); Run-out period end is the latter of seven days after the start of run-out period or the Visit 6 date.

Incidence of marked ECG abnormalities during the placebo run-out period were [redacted] in both treatment groups ([redacted]) (Table 43) (97).

Table 43: Marked ECG abnormalities during treatment withdrawal (97)

ECG parameter Statistic	Daridorexant 50 mg N = 93 n / Nn (%)	Placebo N = 78 N / Nn (%)
ECG Mean Heart Rate (beats/min)		
< 50	[redacted]	[redacted]
> 10 and ≤ 20 decrease from baseline	[redacted]	[redacted]
> 20 decrease from baseline	[redacted]	[redacted]
PR Interval, Single Beat (msec)		
>200	[redacted]	[redacted]
QRS Duration, Single Beat (msec)		

ECG parameter Statistic	Daridorexant 50 mg N = 93 n / Nn (%)	Placebo N = 78 N / Nn (%)
>110	[redacted]	[redacted]
QTcB Interval, Single Beat (msec)		
>450 and ≤ 480	[redacted]	[redacted]
> 30 and ≤ 60 increase from baseline	[redacted]	[redacted]
QTcF Interval, Single Beat (msec)		
>450 and ≤ 480	[redacted]	[redacted]
> 30 and ≤ 60 increase from baseline	[redacted]	[redacted]

ECG=electrocardiogram; HR=heart rate; QTcB=QT interval corrected with Bazett's formula; QTcF=QT interval corrected with Fridericia's formula.
Nn is the number of subjects at risk: those having at least one post-baseline value per ECG parameter for criterion based on post-baseline values only, or those having a baseline value and at least one post-baseline value per ECG parameter for criterion based on change from baseline.

Drug abuse potential

AEs related to drug abuse, dependence and withdrawal were also studied in subjects of study 303. TEAEs suggestive of drug abuse potential were reported for [redacted] in 50 mg group and [redacted] in the placebo group. Intentional overdose was reported by [redacted] in the daridorexant 50 mg group. Accidental overdose was reported for [redacted] in the daridorexant 50 mg group, and [redacted] the placebo group. In addition, [redacted] of accidental overdose were [redacted] (Appendix F) (97).

Sheehan disability scale©

Overall, the SDS© scores showed [redacted] on any of the assessed sub-scores related to daridorexant 50 mg and [redacted] across both treatment groups (97). In both treatment groups, [redacted] from baseline to week 14 [redacted], week 27 [redacted], week 40 [redacted] and placebo run-out period [redacted] (Table 44) (97).

Table 44: Sheehan Disability Scale© – Observed value and change from baseline in total score at week 14, week 27, week 40 and placebo run-out (97)

Time point Statistic	Daridorexant 50 mg (N = 137)			Placebo (N = 128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Week 14						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 27						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 40						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Run-out						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SD=Standard Deviation.

The number of days lost in the week prior to assessment [redacted] daridorexant 50 mg over placebo across all timepoints (week 14, week 27, week 40 and placebo run-out) (97). The [redacted] in days lost was observed at week 14 and week 40 [redacted] (97). The [redacted] for the number of unproductive days, with the [redacted] reported at week 40 [redacted] (Table 45 and Table 46) (97).

Table 45: Sheehan Disability Scale© – Observed value and change from baseline in number of days lost at week 14, week 27, week 40 and placebo run-out (97)

Time point Statistic	Daridorexant 50 mg (N = 137)			Placebo (N = 128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Week 14						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 27						

Time point Statistic	Daridorexant 50 mg (N = 137)			Placebo (N = 128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 40						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Run-out						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SD=Standard Deviation.

Table 46: Sheehan Disability Scale© – Observed value and change from baseline in number of unproductive days at week 14, week 27, week 40 and placebo run-out (97)

Time point Statistic	Daridorexant 50 mg (N = 137)			Placebo (N = 128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Baseline						
n	[redacted]			[redacted]		
Mean (SD)	[redacted]			[redacted]		
Week 14						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 27						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 40						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Run-out						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SD=Standard Deviation.

Additional safety assessments of vital signs (mean of the 2 PSG nights in systolic and diastolic BP and pulse rate), change in body weight, marked laboratory abnormalities, occurrence of suicidal ideation and/or behaviour, next-day residual effects, and rebound insomnia are reported in Appendix F.

B.2.9 Additional analysis of ISI©

B.2.9.1 Seemingly unrelated regression of ISI© scores in study 301

Additional analysis of the ISI© scores in studies 301 and 303 was performed to generate the parameters required for the cost-effectiveness model in B.3.2 Economic analysis. The analysis of ISI© scores in confirmatory study 301 used the seemingly unrelated regression procedure to model the relationship between ISI© scores at month 1 and month 3. Seemingly unrelated regression was used instead of linear mixed-effects model for two reasons. First, it preserves the distinction between the two time points (i.e., month 1 and month 3) ensuring that the regression results exactly predict the observed data points, whereas a linear mixed-effects model would average the treatment effect over the two time points. Second, although seemingly unrelated, the correlation structure between the regressions for each time point is captured and a joint covariance matrix provided for all coefficients which provides the necessary information for the PSA described in B.3 Cost effectiveness.

A total of 557 subjects (out of 620 in the full analysis set) with complete ISI© scores at baseline, month 1 and month 3 were included. Table 47 summarizes the demographic characteristics of subjects included in the analysis. Despite the missing information the characteristics remain similar to the full analysis set (Table 11).

Table 47: Demographic characteristics of subjects in study 301 included in the additional analysis of ISI© scores

Variable Statistic	Daridorexant 50 mg N = 279	Placebo N = 278
Age at screening (years)		
Mean (SD)	56 (15)	56 (15)
Sex [n(%)]		
Male	99 (35%)	89 (32%)
Female	180 (65%)	189 (68%)
Race [n(%)]		
Black or African American	23 (8%)	24 (9%)
Asian	3 (1%)	1 (0%)
White	251 (90%)	251 (90%)
Other	2 (1%)	2 (1%)
Body mass Index (kg/m^2) at Screening		
Mean (SD)	26 (4.3)	26 (4.1)
Region [n(%)]		
US	85 (30%)	96 (35%)
Other (non-US)	194 (70%)	182 (65%)

BMI=Body mass index; SD=standard deviation; US=United States

Table 48 shows the results of the seemingly unrelated regression with ISI© total score as dependent variable and treatment as explanatory variable. After adjusting for baseline ISI© scores, daridorexant 50 mg significantly improved (reduced) ISI© scores compared to placebo at month 1 (-1.70 [95% confidence interval -2.51 to -0.88], p<0.0001) and month 3 (-1.98 [-2.94 to -1.02], p<0.0001).

Table 48: Seemingly unrelated regression of ISI© scores at month 1 and month 3, adjusting for baseline ISI© score

Variable	ISI© score relative to baseline	95% confidence interval	p-value
Month 1			
Daridorexant 50 mg (Reference: placebo)	-1.70	-2.51 to -0.88	<0.0001
Month 3			
Daridorexant 50mg (Reference: placebo)	-1.98	-2.94 to -1.02	<0.0001

ISI©=Insomnia severity index.

Interaction effect between severity of insomnia disorder and treatment on ISI© scores was assessed using the same seemingly unrelated regression model. The results in Table 49 show that while the main effect indicator for the severe subgroup was statistically significant at month 1 (B=2.01 [0.38 to 3.64], p=0.015) and month 3 (B=2.28 [0.35 to 4.21], p=0.021), the severe subgroup indicator by treatment interaction term was not significant at both timepoints (B=0.19 [-1.59 to 1.97], p=0.834 [month 1]; B=-1.44 [-3.55 to 0.67], p=0.18 [month 3]).

Table 49: Seemingly unrelated regression of ISI© scores at month 1 and month 3, stratified by severity of insomnia disorder and adjusting for baseline ISI© score

Variable	Coefficient	95% confidence interval	p-value
Month 1			
Daridorexant 50 mg (Reference: placebo)	-1.72	-2.68 to -0.77	<0.0001
Severe insomnia disorder (Reference: non-severe)	2.01	0.38 to 3.64	0.015
Daridorexant 50 mg * severe insomnia disorder	0.19	-1.59 to 1.97	0.834
Month 3			
Daridorexant 50 mg (Reference: placebo)	-1.54	-2.68 to 0.41	0.008
Severe insomnia disorder (Reference: non-severe)	2.28	0.45 to 4.21	0.021

Variable	Coefficient	95% confidence interval	p-value
Daridorexant 50 mg * severe insomnia disorder	-1.44	-3.55 to 0.67	0.180

ISI©=Insomnia severity index.

B.2.9.2 Attrition in study 303 was associated with smaller change in ISI© scores from baseline

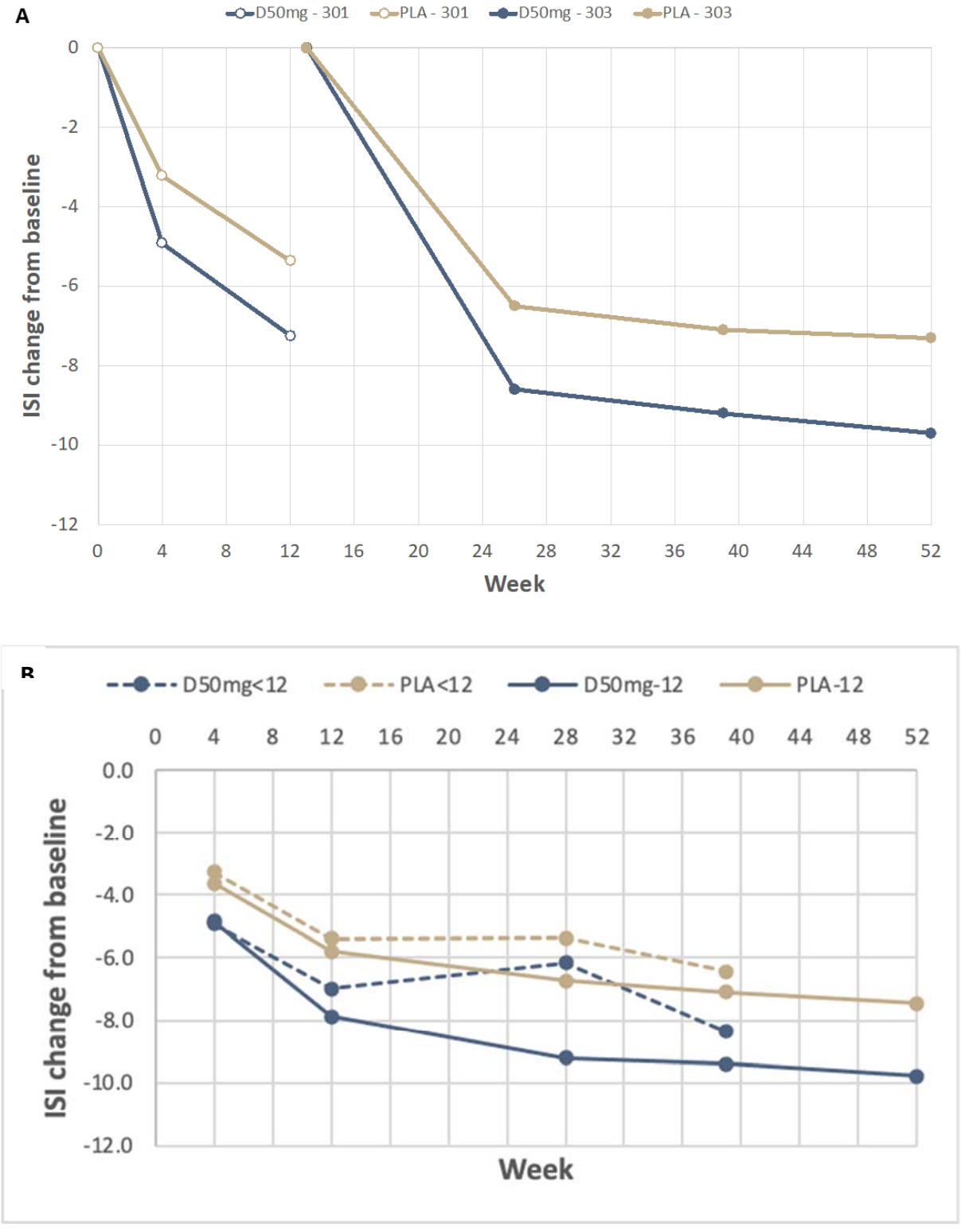
Extension study 303 showed that after the washout period from confirmatory study 301, ISI© scores returned to baseline and that treatment effect was restored upon re-initiation of treatment (Table 36).

The ISI© scores of subjects in the daridorexant 50 mg and placebo groups in confirmatory study 301 and extension study 303 were categorized into two cohorts:

- Subjects who completed the full 12 months of treatment transiting from confirmatory study 301 into extension study 303, and
- Subjects who dropped out of extension study 303 before the week 40 visit.

Figure 13 illustrates the change in ISI© scores from baseline to the end of extension study 303. In both treatment groups, subjects who dropped out of extension study 303 before the week 40 visit had smaller changes in ISI© scores compared to those who completed the study. Visual inspection of the week 28, week 39 and week 52 change scores of subjects who completed the study showed a plateau after week 28. Therefore, the increasing improvement in ISI© scores over time observed in extension study 303 (Table 36) could be attributed to selective attrition.

Figure 13: Change in ISI© scores from baseline to the end of extension study 303, (A) all subjects included and (B) stratified by study completion status



D50mg=daridorexant 50mg, PLA=placebo, D50mg<12=on daridorexant for less than 12 months; D50mg-12=on daridorexant for 12 months; PLA<12=on placebo for less than 12 months; PLA-12=on placebo for 12 months; ISI©=Insomnia severity index.

B.2.10 Quality assessment of the relevant clinical effectiveness evidence

A summary of the quality assessment for studies 301 and 303 is shown in Table 50.

Table 50: Quality assessment of studies 301 and 303

Quality assessment criteria	Grade (Yes/No/Not clear/NA)	
	Study 301	Study 303
Was randomisation carried out appropriately?	Yes	NA
Was the concealment of treatment allocation adequate?	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes
Were the care providers, participants, and outcome assessors blind to the treatment allocation?	Yes	Yes
Were there any unexpected imbalances in dropouts between groups? If so, were they explained or adjusted for?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes

NA=Not applicable.

B.2.11 Meta-analysis

No meta-analysis was undertaken as the 50 mg dose of daridorexant was investigated only in one study (study 301), for which a safety extension study (study 303) was also conducted.

B.2.12 Indirect and mixed treatment comparisons

Not applicable.

B.2.13 Ongoing studies

No ongoing studies.

B.2.14 Innovation

The burden of insomnia disorder is severe, affecting 7.3% of the adult population in England (47). Patients suffering from insomnia disorder experience significant daytime functioning impairment, which negatively impacts physical and mental health, cognitive ability, mood, relationships, QoL and work productivity. The current first-line treatment, CBTi is refused by, or inaccessible to, up to [redacted] of patients. Among those

who do receive CBTi, [redacted] fail to achieve the desired results (1). Current pharmacotherapies are recommended only for short-term use (i.e., <4 weeks, or <13 weeks for melatonin) due to safety concerns and risk of tolerance or dependence. Considering the pervasive nature of insomnia disorder, it often requires longer-term treatments. Therefore, most primary care clinicians prescribe the current pharmacotherapies for longer than their recommended duration due to the lack of alternatives (76, 77).

Daridorexant will be the first DORA to be approved for treatment of insomnia disorder in the European countries. This is an important advancement in the treatment armamentarium of insomnia disorder, in decades. As elaborated in B.2 Clinical effectiveness, results of the phase III trials of daridorexant demonstrated improved night-time and daytime symptoms of insomnia disorder in both adults and the elderly compared to placebo:

- Daridorexant significantly improves objective sleep onset (LPS), sleep maintenance (WASO) and self-reported sleep quantity (sTST).
- Daridorexant improves daytime functioning, as assessed by the patient-reported IDSIQ scores, including total score, sleepiness domain score, mood domain score and alert/cognition domain score. The improved daytime functioning was also seen in work and activity productivity based on Sheehan Disability Score.
- Daridorexant is well tolerated and exhibits an excellent safety profile with no next morning residual effects or withdrawal symptoms upon discontinuation.
- The effects of daridorexant on night-time symptoms and daytime functioning are sustained for up to one year and the safety profile maintained during long-term treatment.

Having daridorexant recommended for use by NICE in primary care has several benefits:

- Access to an efficacious, and safe treatment alternative for patients with insomnia disorder, addressing the different facets of the condition in a sustainable manner, including night-time and daytime symptoms with limited adverse events, no

rebound, no withdrawal symptoms upon discontinuation, no tachyphylaxis, and most importantly, no next morning sleepiness. Treatment with daridorexant, which can be prescribed long-term is expected to improve patients' QoL and potentially improve work productivity.

- Minimizes use of inappropriate current pharmacotherapies beyond their recommended duration and reduces off-label use of pharmacotherapies (e.g., sedating anti-depressants).

- A safer alternative for elderly patients with comorbidities or adults receiving treatments that make them unsuitable for existing pharmacotherapies.

B.2.15 Interpretation of clinical effectiveness and safety evidence

The results from phase III confirmatory study 301 and extension study 303 demonstrated that treatment with daridorexant 50 mg once daily in patients with insomnia disorder was superior to placebo for the primary endpoints of objective sleep induction (LPS) and maintenance (WASO). Its superiority over placebo was also demonstrated for key secondary subjective endpoints of patient reported sleep quantity (sTST) and daytime functioning (IDSIQ scores) (96). The benefits of daridorexant on patient-reported sleep quantity and daytime functioning were sustained for up to a year, as shown by the improved sTST and IDSIQ sleepiness domain score at month 6, month 9 and month 12 (96).

Treatment with daridorexant 50 mg once daily led to a clinically and statistically significant reduction in WASO by 22.8 minutes and LPS by 11.4 minutes at month 1 compared with placebo. The results met the pre-specified effect size of 0.37, which is comparable with those of other pharmacotherapies for insomnia disorder (101). The reductions were consistent throughout the confirmatory study 301, with daridorexant 50 mg significantly reducing WASO by 18.3 minutes and LPS by 11.7 minutes at month 3 compared with placebo. The treatment effect with daridorexant became apparent shortly after randomisation and was maintained throughout the trial (96). The results were consistent across the pre-defined subgroups stratified by age, gender and geographical region as indicated by the overlapping confidence intervals of the primary and key secondary efficacy endpoints (Figure 13).

Consistent with improvements in objective sleep onset and maintenance, treatment with daridorexant 50 mg once daily demonstrated statistically significant improvements in patient-reported sTST and IDSIQ sleepiness domain score. Self-reported TST improved by 22.1 minutes and 19.8 minutes at month 1 and month 3, respectively compared with placebo (96). Similarly, IDSIQ sleepiness domain score reduced by 1.75 and 1.90 points at month 1 and month 3, respectively compared with placebo. Overall, confirmatory study 301 demonstrated statistically significant improvements in objective and subjective sleep measures among patients treated with daridorexant 50 mg once daily compared to those treated with placebo (96).

The results of extension study 303 showed that the effects of daridorexant on patient-reported sleep quantity and quality were [redacted] for up to a year on treatment (97). Compared with placebo, daridorexant 50 mg once daily led to [redacted] in sTST of [redacted] at month 6, but [redacted] of [redacted] at month 9 and month 12, respectively (97). IDSIQ sleepiness domain scores [redacted] across all timepoints among patients treated with daridorexant compared to those treated with placebo (97).

A major challenge in patients with insomnia disorder is to reverse impaired daytime functioning (102). As discussed in B.1.3.3 Humanistic burden, most of the existing pharmacotherapies can further deteriorate daytime functioning. Subjects in the daridorexant 50 mg group reported improvements in all aspects of daytime functioning compared to placebo in [redacted] confirmatory study 301 [redacted], as assessed by IDSIQ total score, mood domain score, alert/cognition domain score, and sleepiness domain score (96). In addition, responder analysis of the IDSIQ sleepiness domain score in confirmatory study 301 showed a higher proportion of subjects achieving a 4-point reduction (clinical change threshold) in the daridorexant 50 mg group than in the placebo group (35). The ability to reverse impaired daytime functioning demonstrated by daridorexant addresses an important gap in the current insomnia disorder treatment landscape.

Analysis of ISI© scores in studies 301 [redacted] provided additional evidence to support the superiority of daridorexant over placebo on the primary and key secondary efficacy endpoints. The ISI© scores of subjects treated with daridorexant 50 mg reduced significantly from baseline by 1.7 and 2.0 at month 1 and month 3, respectively

compared to subjects treated with placebo. The benefits were consistent across insomnia severity subgroups as indicated by the non-significant interaction between treatment and severity (97, 99).

Placebo effects in insomnia clinical trials are commonly observed, and studies have shown that such effects appear to be robust and durable in longer-term trials (103-106). This was observed in studies 301 and 303, where subjects in the placebo group experienced [redacted] in ISI© scores from baseline. Of note, the [redacted] after 12 weeks of placebo treatment when accounting for selective attrition (Figure 13), indicating [redacted] effect over time. The [redacted] was demonstrated in a meta-analysis of 27 placebo-controlled insomnia RCTs, where [redacted] of placebo effects was achieved after 9 to 12 weeks of treatment (107). This was reflected in the modelling of ISI© scores in the base case cost-effectiveness model presented in B.3 Cost .

In terms of safety, daridorexant was well tolerated in both adult and elderly patients. In confirmatory study 301, the incidence of somnolence was low among subjects in the daridorexant 50 mg group and was even numerically lower than that in the placebo group (Table 18), likely attributed to better sleep at night (99). Nausea, headaches, mild dizziness, and fatigue were slightly more frequent in the daridorexant group than placebo group, whereas the incidence of falls was slightly lower in the daridorexant group than placebo group (Table 18) (99). These findings were consistent with those of extension study 303 (Table 38). In both studies, daridorexant demonstrated no next-morning residual effects, no evidence of abuse potential, and no signs of rebound insomnia or withdrawal symptoms upon treatment discontinuation (97).

The benefits of daridorexant have been demonstrated in both adults and elderly and across a range of insomnia severity. In conclusion, the evidence submitted in the current company submission demonstrated that daridorexant is associated with clinically meaningful improvements in objective sleep onset and maintenance, and subjective sleep quantity and quality as well as daytime functioning. Daridorexant is the first DORA in the European countries demonstrating safety and efficacy for up to one year, with no next-morning residual effects and no risk of rebound insomnia or withdrawal symptoms upon treatment discontinuation. Thus, its recommendation for

the treatment of insomnia disorder in primary care will improve outcomes, increase work productivity, and benefit the patient population treated by the NHS.

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

A SLR was conducted to identify existing economic evaluations for the treatment of insomnia disorder. Full details of the process and methods used are described in Appendix G.

In summary, while the SLR identified studies of other technologies, they were deemed as not suitable comparators to daridorexant. In addition, no studies relevant to daridorexant were identified and therefore no quality assessment was conducted.

B.3.2 Economic analysis

Due to the lack of previously published economic evidence for daridorexant, a *de novo* economic model is included in the submission. The following cost-effectiveness analysis demonstrates the incremental cost effectiveness ratio (ICER) of daridorexant in comparison to no treatment for patients with insomnia disorder.

B.3.2.1 Model structure

The model structure is illustrated as pathways in Figure 14 showing that clinical trial evidence is available for direct estimation of the treatment effect on ISI©, on side effects and on productivity losses (through the SDS©). HCRU, EQ-5D and WPAI impacts are captured indirectly through mapping from ISI© using an external data source.

Treatment has a direct cost of [redacted] per day. The effectiveness of treatment is captured by its impact on the ISI©, a PRO that was included in both study 301 and 303 (97, 99). The SDS© was also directly measured in the clinical trial programme but is shown as a dotted line in Figure 14 as its inclusion is a non-reference case analysis.

As EQ-5D was not included in the clinical study, a second data source, the Cerner Enviza NHWS was utilized (B.1.3.2 Epidemiology). The NHWS included subjects that reported insomnia symptoms and who then completed the ISI© alongside EQ-5D. Since the same survey also included questions on direct health care resource use (GP visits, emergency room attendances and inpatient admissions), as well as the WPAI, it was possible to use the same data source to also look at potential cost-savings (both

direct health service costs and indirect productivity losses) associated with reducing insomnia severity. The pathway from ISI© to productivity to cost-effectiveness is shown as a line in Figure 14, as productivity losses are excluded from NICE's reference case of methods and are only included as a scenario analysis.

Shown in gold in Figure 14 is the potential SAE pathway. However, the label only identified headache and somnolence as potential side-effects of treatment and noted that these were not statistically significantly more frequent than placebo based on the registration trials (Table 18 and Table 38). Therefore, SAE was not included in the model.

In line with NICE guidance and following discussions with NICE as part of the decision problem meeting, the timeframe of the model is limited to a short-term time horizon in the base case.

We consider a short-term time horizon to be appropriate for several reasons:

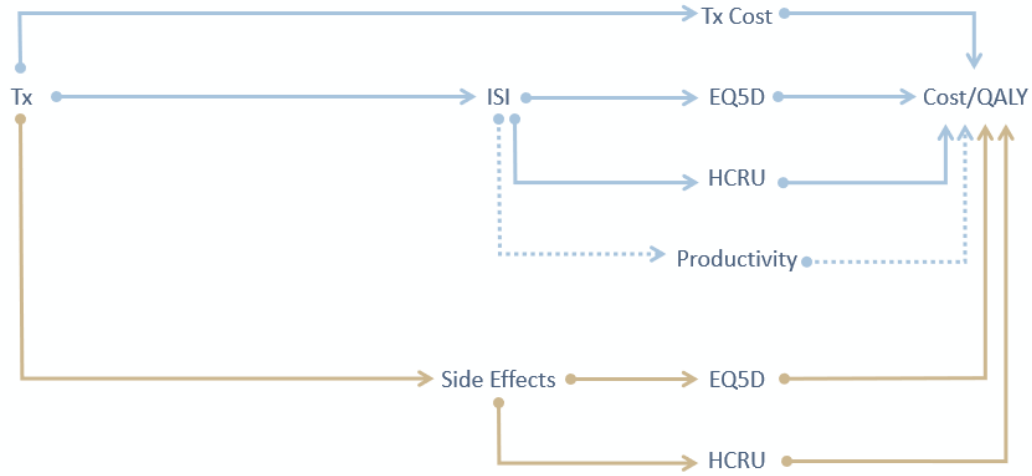
- No mortality effects for insomnia treatment are assumed. Therefore, the only impact is on HRQoL (as measured by the EQ-5D).
- Daridorexant has quick onset and a short half-life. Therefore, the treatment benefit occurs when taking the drug and that treatment effect stops when treatment stops, as demonstrated by the placebo run-out phase in-between study 301 and 303.
- The label for daridorexant suggests that all patients initiating the drug should undergo clinical review at three months.

As the Idorsia clinical trial programme was based on a 12-week confirmatory trial (301) and a 40-week extension study (303), we chose a 12-month time horizon for the model (97, 99). It should be noted, however, that this timescale is for convenience; the model could be presented for as short as a single day, since the assumption is that the benefit is obtained when treatment is taken and lost once treatment stops. Based on extension study 303, we assume that there will be [redacted] among patients who received less treatment benefit as measured by ISI© (97). The assumption is that beyond three

months, no further increase in ISI© occurs for the individual patient, but that drop out occurs among those with less treatment benefit (B.2.9 Additional analysis of ISI©).

Nevertheless, and as pointed out by NICE during the decision problem meeting on 23rd April 2022, better sleep could have long term benefits on overall mortality through reduction of road traffic accidents, reduced cardiovascular stress, and reduction in falls (12). There is some epidemiological evidence for such long-term effects and therefore, as a scenario analysis (B.3.11.3 Scenario analysis), we look at the potential long-term cost-effectiveness of daridorexant when such long-term impacts are factored in.

Figure 14: Pathway from treatment to value via ISI©, EQ-5D and cost



Blue solid lines show main analysis (reference case). Gold solid lines show typical reference case pathway excluded from this evaluation due to lack of any serious safety concerns apparent in the data. Blue dotted lines show two potential routes to estimating productivity (non-reference case) either directly from the trial (SDS©) or mapped from ISI© (NHWS).
Tx=treatment; ISI=Insomnia Severity Index; HCRU=healthcare resource utilisation; QALY=quality-adjusted life year

As highlighted in B.1.3.5 Clinical care pathway, evidence for long-term treatment of insomnia disorder is lacking and existing guidelines are directed towards short-term treatment options. Similarly, previous evaluations by NICE, including TA77 and MTG70, focused on technologies indicated for short-term management of insomnia symptoms and not insomnia disorder (74, 108). They are therefore not comparable with daridorexant's intended population. Nevertheless, a comparison of the main inputs of the economic models between the past evaluations and this submission are summarised in Table 51.

Table 51: Comparison of main inputs of economic model between previous and current evaluations

	Previous evaluations		Current evaluation	
Factor	**TA77 (108)**	**MTG70 (74)**	**Chosen values**	**Justification**
Time horizon	Short-term	Short-term	12 months	As presented in B.3.2.1 Model structure
Treatment waning effect?	Not included	Not included	Not included	Short-term model based on observed data for base case so waning not relevant. No waning included in lifetime model due to lack of evidence of waning.
Source of utilities	Clinical trial	Clinical trial	Mapping of ISI© scores to EQ-5D	EQ-5D was not collected in daridorexant's clinical trial programme
Source of costs	Publicly available information	Publicly available information	Publicly available information	In line with NICE recommendations

ISI=Insomnia Severity Index

B.3.2.2 Intervention technology and comparators

The intervention technology, daridorexant, is a DORA that blocks the binding of neuropeptides orexins, which is connected to sleep/ wake regulation. This model considers patients utilizing a 50 mg dosage administered once per day.

Current treatment guidelines focus only on short-term treatment options, and none include recommendations for long-term treatment of insomnia disorder (B.1.3.5 Clinical care pathway). The NICE CKS suggests non-pharmacological therapy (sleep hygiene and face-to-face or digital CBTi) as first line treatment. For severe cases that do not respond to non-pharmacological therapy, a short course of hypnotic drug treatment is proposed, but only for one week. Melatonin can be used, but only for up to 13 weeks and in adults aged ≥55 years. Daridorexant is the first insomnia treatment with longer term data for the treatment of insomnia disorder. For this reason, the no treatment comparator is appropriate (as per the final scope), and the placebo arm of the trial serves as a proxy for no treatment based on the analysis of study 301. The intention-to-treat treatment effect for daridorexant that is used in the model is adjusted for the placebo comparison for the first 3 months.

B.3.3 Clinical parameters and variables

B.3.3.1 ISI© as the key clinical parameter

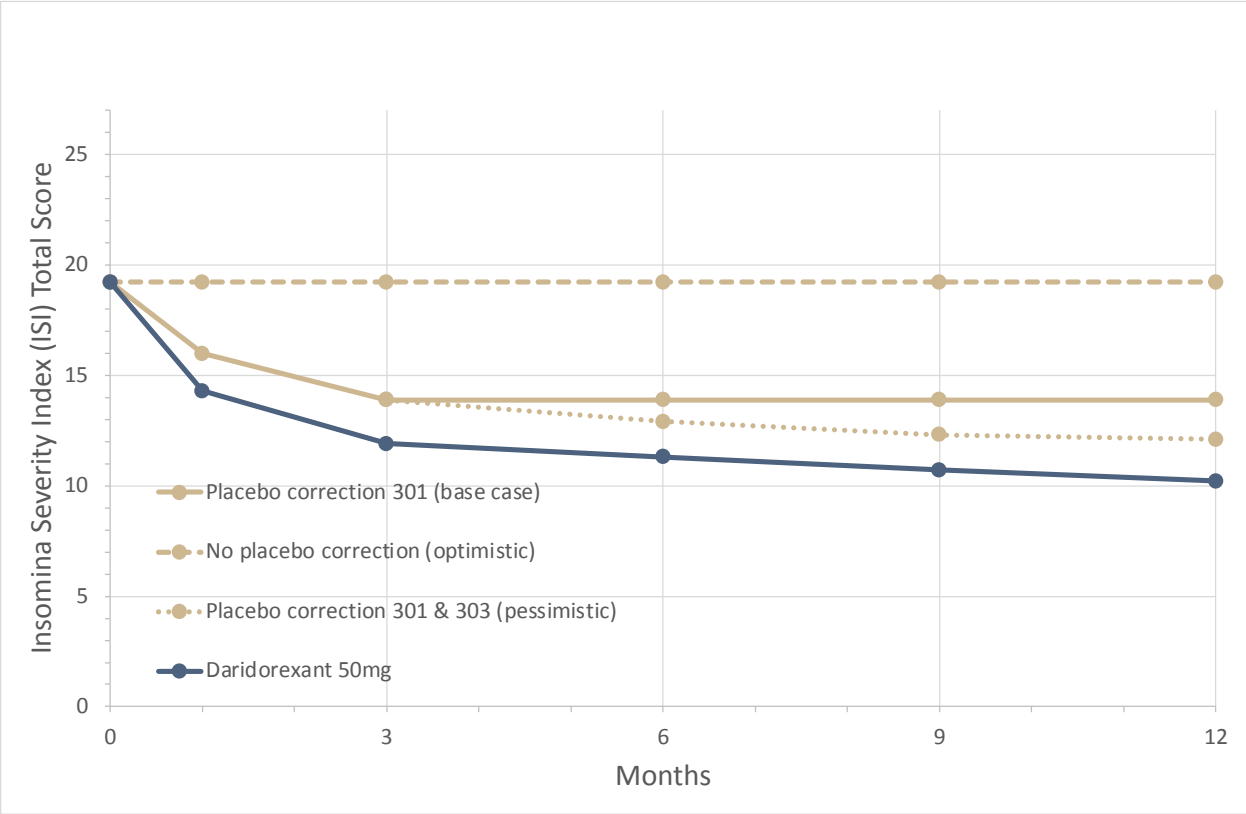
The clinical outcome driving the model is the ISI©, which was an exploratory endpoint in the daridorexant phase III clinical trials (97, 99). The ISI is described in detail in B.1.3.1 Disease overview. Briefly, the ISI© assesses insomnia based on criteria from the ICSD-3 and is currently one of the most used insomnia-specific PRO questionnaires. It has already been translated into more than 50 languages and has been validated as a treatment response metric for insomnia patients (37, 109, 110). Continuous ISI© scores were utilized in the modelled population to derive EQ-5D scores that informed QALY calculations. Other clinical endpoints assessed in the trial include WASO, LPS, sTST and IDSIQ scores; however, these outcomes were not utilized in the model because there was no mechanism to link these outcomes to EQ-5D.

B.3.3.2 Base case plus best/ worst case scenarios of ISI© trajectory

In Figure 15 below the modelled trajectory of ISI© is presented based on the analysis of study 301 and 303. The analysis of study 301 was conducted using the seemingly unrelated regression procedure as reported in B.2.9 Additional analysis of ISI©. The use of seemingly unrelated regression allowed for the usual estimation of treatment effect that adjusts for both baseline ISI© and placebo, while further allowing the correlation between month 1 and month 3 observations to be captured for use in the PSA. For extension study 303, the assumption was that the improvements in ISI© over time were due not to an increasing effect of treatment, but to selective attrition (as argued in B.2.9 Additional analysis of ISI© where the lower ISI© change from baseline for patients dropping out of 303 was presented in Figure 13). We therefore modelled treatment discontinuation based on the observed discontinuation rates in both studies in the treatment group only. For the no treatment group, we modelled placebo adjustment based on study 301 only, since study 303 again presented evidence of selective attrition (see Figure 13), whereas in practice (and in our model), patients are unable to dropout from 'no treatment'. Our base case assumption was that after study 301, no treatment patients would continue at the same ISI© achieved by the end of study 301 (plain line in Figure 15). Of course, in practice, patients without treatment would not be expected to gain anything relative to their stable baseline ISI©. Therefore,

we model as a more optimistic scenario, that the full change from baseline score is attributable to treatment. This is shown in Figure 15 as the upper dashed line for no treatment ISI©. A more pessimistic scenario would be to continue to placebo adjust into the period of study 303 and this is shown by the lower dotted line for no treatment ISI. It is clear from Figure 15 that the chosen base case is toward the more pessimistic and represents a conservative estimate of the true value of daridorexant.

Figure 15: Modelled trajectory of ISI© from phase III study 301 and 303 extension study showing base case, optimistic and pessimistic scenarios regarding placebo adjustment



The discontinuation rates are [redacted] in study 303 which is a [redacted] clinical practice than in study 301 (97, 99). Nevertheless, until daridorexant is approved, it is not possible to observe real-world treatment discontinuation. As part of the deterministic sensitivity analysis (DSA) we present a simultaneous low and high value of the discontinuation rates shown in Figure 16 using the lower and upper confidence limits.

When incorporating the impact of discontinuation into the cost-effectiveness analysis using the following two assumptions that are conservative with respect to the estimated cost-effectiveness of daridorexant.

1. That discontinuation occurs at the midpoint of the periods such that the estimated QALYs for the period are based on the average of the EQ-5D at the start and at the end of the period.

2. That cost of treatment is incurred for the full period assuming prescriptions are filled at the start of the period before discontinuation occurs.

Figure 16: Rate of treatment discontinuation in study 301 (months 1 and 3) and study 303 (months 6, 9 and 12)

B.3.4 Measurement and valuation of health effects

B.3.4.1 HRQoL data from clinical trials

Condition-specific HRQoL data were captured in both study 301 and 303 in the form of the ISI© (97, 99). The analysis of these data is described in B.2.9 Additional analysis of ISI©, and the model based on these data is outlined in B.3.3 Clinical parameters and variables. Since no EQ-5D data were available in the clinical studies, a mapping exercise was conducted based on the NHWS dataset where both the EQ-5D and the ISI© instruments were collected for cohorts self-reporting insomnia. The NHWS dataset is described in detail in B.1.3.2 Epidemiology.

B.3.4.2 Mapping

Health-related quality-of-life studies

As discussed in B.1.3.3 Humanistic burden, the SLR did not identify HRQoL studies relevant for the model presented in this submission. Therefore, a novel mapping algorithm was used to derive EQ-5D utilities from ISI© scores reported in study 301 and 303 as described below.

A generalised linear model (GLM) was used to create the mapping function from the cross-sectional NHWS survey (40). Since EQ-5D is bounded at 1 from above and has a left skew distribution (see Figure 17), we employed a standard linear transformation to the disutility (dU) scale with dU = 1 – U, where U is the observed EQ-5D utility, such that disutility was bounded from below at zero and the distribution is right skewed.

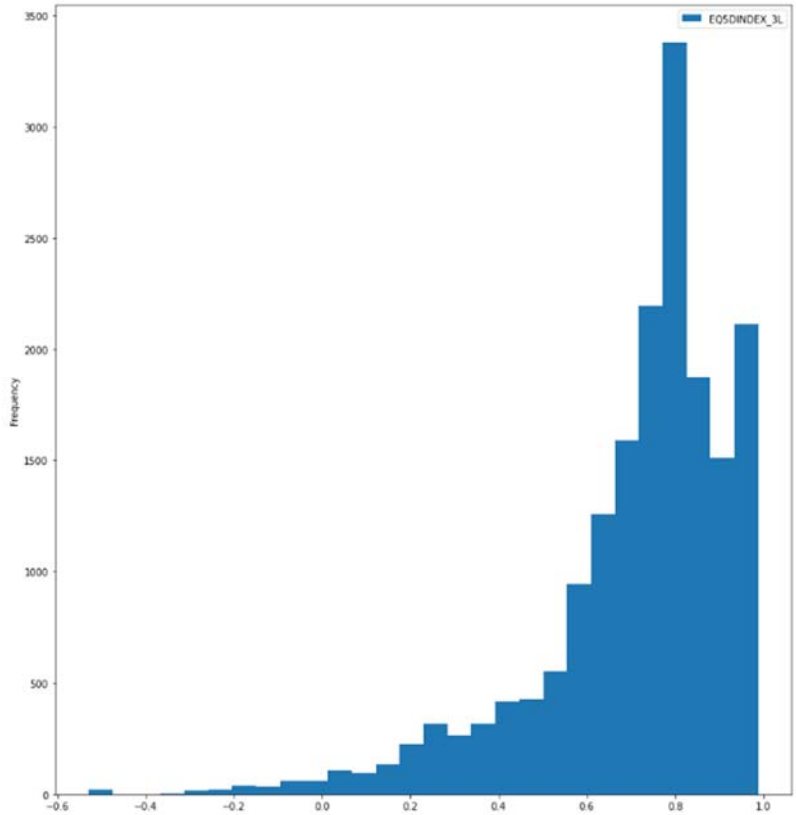
The disutility then formed the explanatory variable in a GLM with a gamma distribution family and log link function such that:

$$E[dU_i] = exp\left\{\alpha + \sum_{j=1}^{k} \beta_j X_{ij} + \gamma ISI_i\right\}$$

where the β_j's represent coefficients *j* on *k* explanatory variables X_j and where γ is the estimated coefficient on the observed ISI© score. Employing the simple linear retransformation back to the utility scale yields the following mapped EQ-5D utility:

$$E[U_i] = 1 - exp\left\{\alpha + \sum_{j=1}^{k} \beta_j X_{ij} + \gamma ISI_i\right\}.$$

Figure 17: Distribution of the EQ-5D tariff scores in the NHWS dataset



EQ-5D=EuroQol-5D; NHWS=Cerner Enviza National Health and Wellness Survey

Adverse reactions

The inclusion of TEAEs was explored. However, as shown in B.2 Clinical effectiveness (Table 18 and Table 38) we see little difference in AE rates between the daridorexant 50mg and placebo groups, indicating a favourable safety profile of daridorexant. The AEs that were reported in ≥2% in either treatment groups included nasopharyngitis, headache, fatigue, dizziness and nausea. Since these side effects were reported to be mild, they are expected to have a negligible impact on HRQoL and patient costs and were excluded from the model.

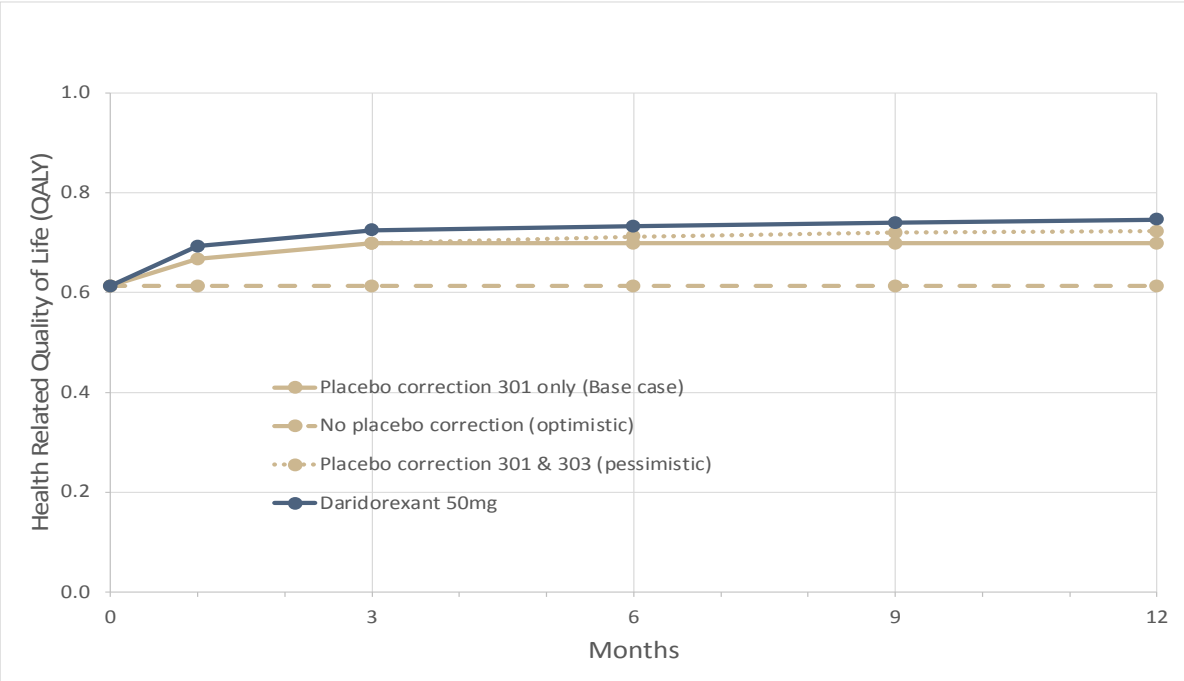
Health-related quality-of-life data used in the cost-effectiveness analysis

The fitted mapping function based on NHWS data, as described in B.3.4.1 HRQoL data from clinical trials, generated a model as follows (40):

$$E[U] = 1 - exp\{-1.849 + 0.047 \cdot ISI\}.$$

Applying this mapping equation to the observed ISI scores making up the profile presented in Figure 15 yields the profile of HRQoL utility as shown in Figure 18, including the optimistic and pessimistic scenarios for ISI discussed above.

Figure 18: Health Related Quality of Life utility profile of EQ-5D mapped from ISI©



EQ-5D=EuroQol-5D

Taking the predicted utility at the end of study 301 and at the end of 12 months (after adjusting for selective dropout), the predicted EQ-5D utilities in the model and the associate utility gains (incremental utility) are presented in Table 52.

Table 52: Utility values for ISI© score at different timepoints

State	Utility value: mean (SE)	(95% CI)	Justification
No Treatment* (ISI© = 13.9)	0.698 (0.008)	(0.687-0.718)	Mapping from ISI© described above
Daridorexant* (ISI© = 11.9)	0.725 (0.007)	(0.712-0.737)	Mapping from ISI© described above
Daridorexant** (ISI© = 9.4)	0.746 (0.010)	(0.722-0.760)	Mapping from ISI© described above
Incremental utility*	0.027 (0.006)	(0.011-0.033)	Mapping from ISI© described above
Incremental utility**	0.048 (0.010)	(0.020-0.056)	Mapping from ISI© described above

*At the end of study 301
**For those remaining on treatment at the end of 12 months
ISI=Insomnia severity index

B.3.5 Cost and healthcare resource use identification, measurement, and valuation

B.3.5.1 Intervention and comparators' costs and resource use

As per the NICE scope, the relevant treatment is the absence of pharmacotherapy. Therefore, there are no treatment costs for the no treatment group. The daily cost of daridorexant is anticipated to be [redacted] per day, giving an annual cost of [redacted] as shown in Table 53. In the cost-effectiveness model, we also included treatment discontinuation using the assumptions identified above. Therefore, Table 53 also includes the 12-month cost of daridorexant after accounting for treatment discontinuation.

Table 53: Intervention and comparator costs

Drug	Cost per day	Annual cost
Daridorexant 50mg	[redacted]	[redacted]
Daridorexant 50mg*	NA	[redacted]
No Treatment	£0	£0

*Discontinuation-adjusted 12-month cost
NA=not applicable

B.3.5.2 Health-state unit costs and resource use

The association between direct health care resource use and ISI© were calculated from the NHWS using a GLM with a negative binomial distribution family and a log link. Table 54 gives the unit costs for each category of resource use that were attached to the predicted health care resource use of each type.

Table 54: Unit costs for health care resource use

Resource	Unit Cost (2021)	Source
General practitioner	£39.23	PSSRU 2021 (111)
Emergency room	£184.62	NHS England 2019/20 (112)
Inpatient	£996.29	NHS England 2019/20 (112)

PSSRU=Personal Social Services Research Unit; NHS=National Health Service

In Figure 3 the predicted direct health care cost of each category of resource use is shown as a stacked bar chart for each point of the ISI© scale, where the value shown relates to the reduction from one-point higher on the ISI© scale. These are the values

that are used in terms of cost-offsets in the model of the corresponding improvement in ISI© associated with treatment.

B.3.5.3 Adverse reaction unit costs and resource use

As described above, the AEs reported in study 301 [redacted] were mild not expected to have significant impact on HRQoL and patient costs. This was reflected in the regulatory label of daridorexant (3) and consequently no AEs were included in the model.

B.3.5.4 Miscellaneous unit costs and resource use

Although not part of the reference case, productivity losses are an extremely important part of the impact of insomnia disorder. As described in B.3.2 Economic analysis, the model structure allows for two alternative ways of estimating productivity losses from chronic insomnia disorder: 1) directly from the SDS© included in the clinical programme, and 2) indirectly from the WPAI mapped to ISI in the NHWS dataset. Each approach is described in detail below.

<u>Directly estimated from the clinical trial programme (SDS©)</u>

As described in B.1.3.1 Disease overview, the SDS© was collected on both study 301 and 303 (97, 99). To estimate productivity losses for absenteeism in the model, the hourly median wage rate of £13.58 (113) was applied to the level of absenteeism assuming 4.90 working days per week and 7.5 hours for each working day inflated to the relevant time period (number of weeks) in the model.

The same method was applied to the whole time-equivalent (WTE) days lost to presenteeism. The calculation of WTE was based on weighting the days where patients reported they were underproductive (question 2 of the SDS©) by the level of unproductiveness (Likert scale 1 of the SDS©) converted from a 10-point scale to a percentage.

The results for the 12-month period of the model are presented in Table 55 and show that over the first year of the model, for a patient remaining on treatment for the full year, productivity savings amount to [redacted] more than the annual cost of treatment.

Table 55: Productivity losses estimated directly from Sheehan Disability Scale© results in study 301 and 303

	Baseline*	Month 1	Month 3	Month 6	Month 9	Month 12	TOTAL
Placebo							
Absenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Presenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Daridorexant							
Absenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Presenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference							
Absenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Presenteeism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Baseline is included and calculated for preceding month for comparison with month 1 but is not included in the total column

Indirectly estimated by mapping WPAI to ISI© in the NHWS

The NHWS dataset included administration of the WPAI. This afforded the opportunity to examine the indirect effect of insomnia on work productivity. The percentage absenteeism and presenteeism were estimated from the WPAI using the standard algorithm. These then formed the explanatory variable in a log-link GLM with ISI© as an explanatory variable. Percentage absenteeism as a function of ISI© was then costed utilising the median annual wage rate of £25,971 (113). Percentage presenteeism was applied as a weighting only to the percentage of time that subjects were present at work (that is to the 1 – percentage absenteeism) and was also costed using the median annual wage rate. Combining these two estimates together allowed the total productivity losses associated with insomnia disorder to be estimated as a function of ISI© score as presented earlier in Figure 4.

Note that the negative presenteeism figures shown in Figure 4 arise due to the relationship between absenteeism and presenteeism. At high values of ISI©, the high rates of absenteeism reduce presenteeism with corresponding increases in presenteeism if ISI© score is reduced.

B.3.6 Severity

Not relevant to this submission.

B.3.7 Uncertainty

Not relevant to this submission.

B.3.8 Managed access proposal

None at this time.

B.3.9 Summary of base-case analysis inputs and assumptions

B.3.9.1 Summary of base-case analysis inputs

All model parameters for the base case model are listed in Table 56. Additional parameters to extend the model to a lifetime horizon are presented in B.3.11.3 Scenario analysis.

Table 56: Full list of parameters for the model with names, values, description and the distribution used for the probabilistic analysis

Name	Value	Prob Dist	Description
ISI© parameters at baseline			
baseISI	19.21	normal	Baseline ISI© score at model start for whole population (group 1 when applicable)
baseISI2	0.00	normal	Baseline ISI© score at model start for group 2 (when applicable)
Parameters from seemingly unrelated regression for ISI© at month 1			
consM1	3.94	multinormal	Constant from SUREG analysis: month 1
baseISIadjM1	0.63	multinormal	Baseline adjustment coefficient from SUREG analysis: month 1
D50mgM1	-1.70	multinormal	Treatment effect of Daridorexant 50mg from SUREG analysis: month 1
sgMEM1	0.00	multinormal	Main effect for subgroup (when applicable): month 1
sgTxintM1	0.00	multinormal	Subgroup by treatment interaction (when appropriate): month 1
Parameters from seemingly unrelated regression for ISI© at end of follow up (month 3)			
cons	3.77	multinormal	Constant from SUREG analysis: end of follow-up
baseISIadj	0.53	multinormal	Baseline adjustment coefficient from SUREG analysis: end of follow-up
D50mg	-1.98	multinormal	Treatment effect of Daridorexant 50mg from SUREG analysis: end of follow-up
sgME	0.00	multinormal	Main effect for subgroup (when applicable): end of follow-up
sgTxint	0.00	multinormal	Subgroup by treatment interaction (when appropriate): end of follow-up
ISI© parameters from 303			
303ISI6m	-0.60	normal	303 six-month change from end 301 for whole population (treatment)

Name	Value	Prob Dist	Description
303ISI9m	-1.20	normal	303 nine-month change from end 301 for whole population (treatment)
303ISI12m	-1.70	normal	303 twelve-month change from end 301 for whole population (treatment)
303ISI6mPLA	0.00	normal	303 six-month change from end 301 for whole population (placebo)
303ISI9mPLA	0.00	normal	303 nine-month change from end 301 for whole population (placebo)
303ISI12mPLA	0.00	normal	303 twelve-month change from end 301 for whole population (placebo)
Drop out / persistence			
dom1	0.04	beta	Drop out from initiation to month 1
dom3	0.05	beta	Drop out from month 1 to month 3
dom6	0.24	beta	Drop out from month 3 to month 6
dom9	0.09	beta	Drop out from month 6 to month 9
dom12	0.13	beta	Drop out from month 9 to month 12
Utility mapping			
consU	-1.849	multinormal	Constant term for utility from NHWS
ISItotU	0.047	multinormal	Coefficient for utility from NHWS
Direct costs			
TxC	[redacted]	NA	Daily cost of Daridorexant 50mg
consGP	1.88	multinormal	Constant term for general practitioner (GP) visit from NHWS model (output 3)
ISItotGP	0.02	multinormal	Coefficient for GP visit from NHWS model (output 3)
consER	-1.18	multinormal	Constant term for emergency room (ER) visit from NHWS model (output 3)
ISItotER	0.03	multinormal	Coefficient for ER visit from NHWS model (output 3)
consIP	-1.80	multinormal	Constant term for inpatient (IP) visit from NHWS model (output 3)
ISItotIP	0.02	multinormal	Coefficient for IP visit from NHWS model (output 3)
ucGP	£39.23	NA	Cost of a GP visit (PSSRU)
ucER	£184.62	NA	Cost of an ER visit (National Cost Collection for the NHS 2020)
ucIP	£996.29	NA	Cost of an IP visit (National Cost Collection for the NHS 2020)
annualWR	£25,971	NA	Annual wage rate (median) in the UK (ONS Annual Survey of Hours and Earnings 2021)

ISI=Insomnia severity index; SUREG=seemingly unrelated regression; NHWS=Cerner Enviza National Health and Wellness Survey; PSSRU=Personal Social Services Research Unit; NHS=National Health Service; UK=United Kingdom

B.3.9.2 Assumptions

- AE rates and costs are not included in the economic analysis as their impact on the incremental outcomes between daridorexant and placebo are negligible.

- The model time frame is 12 months. We expect this to capture applicable patient outcomes as the link of treatment to long-term effects such as falls, and cardiovascular mortality have not been proven.

- The cost of daridorexant is assumed to be [redacted] per day as per the manufacturer's recommendation.

- Assumed that there were no deaths in the 12-month model timeframe.

B.3.10 Base-case results

B.3.10.1 Base-case incremental cost-effectiveness analysis results

The base case results of the cost-effectiveness model are presented in

Table 57, which shows the estimated costs and QALYs under treatment and no treatment assuming 100% persistence. Also shown are the incremental results once persistence is taken into account. Because the model is conservative and assumes that when subjects stop treatment the full cost of prescribing treatment for the period in question is applied, but full benefits are not accrued due to stopping treatment we see that the persistence adjusted results lead to slightly higher ICERs. Since the 12-month cost-effectiveness is between the thresholds of £20,000 per QALY and £30,000 per QALY, we see a negative net-health-benefit (NHB) at the lower threshold and a positive NHB at the higher threshold.

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034

Technology	Cost	QALY
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

Table 57: Base case cost-effectiveness results for the 12-month model

*Adjusted for persistence
**95% uncertainty intervals from the probabilistic analysis
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Disaggregated results for the costing are provided in Appendix J. This includes event rates for GP visits, emergency room (ER) attendances and inpatient (IP) stays. We also show how the cost-effectiveness of daridorexant 'evolves' over the first 12-month period reflecting the short-term time horizon of the presented model. The results presented in

Table 57 represents an average for the first 12-months of the model. Due to the impact

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

of selective attrition in the model – for those subjects remaining on treatment at the end of 12 months, their estimated cost-effectiveness is [redacted] per QALY going into subsequent years of the model – a considerable improvement compared to the base case result of [redacted] per QALY for the average over the first 12-months. Cost-effectiveness of daridorexant beyond 12 months is detailed in Appendix J and further developed in the lifetime scenario analysis in B.3.11.3 Scenario analysis.

Clinical outcomes for the model have already been presented in B.2 Clinical effectiveness since no extrapolation was necessary.

B.3.11 Exploring uncertainty

A genuine attempt has been made to include all uncertainties into the base case model. Because the short-term model is informed by a robust clinical programme, made up of the study 301 and 303, it has been possible to characterise much of the uncertainty statistically through patient-level data analysis. The parameter values reported in Table 56 have all been included in the PSA. Additional elements of uncertainty, not captured statistically, include:

- The representativeness of the trial-based persistence rates to real-world practice
- The lack of direct measurement of EQ-5D in the clinical trial programme and the reliance on a mapping algorithm to map ISI to EQ-5D

Uncertainty about real-world persistence with treatment can only be resolved once treatment is recommended for use in the NHS. However, the evidence of a quick onset of daridorexant effectiveness and of the corresponding loss of effectiveness once treatment is discontinued mitigates the uncertainty in cost-effectiveness, since the costs and benefits move together.

Despite the lack of direct measurement of EQ-5D in daridorexant's clinical trial programme, the use of the large scale, representative NHWS data has allowed a robust mapping algorithm to be generated and uncertainty related to both statistical estimation and representativeness of the sample has been incorporated.

B.3.11.1 Probabilistic sensitivity analysis

The summary information presented in Table 56 relating to the parameters included in the probabilistic sensitivity analysis (PSA) is supplemented in Appendix N with full details of the parameters of the statistical distributions for all the parameters included in the PSA with the exception of the ISI© estimates. These were obtained from the study 301 and 303, and are reported in B.2 Clinical effectiveness. For the seemingly unrelated regressions, correlations between the regression coefficients are captured using Cholesky decomposition of the associated covariance matrices. The overall cost-effectiveness results for the base case analysis are presented on the cost-effectiveness plane (Figure 19). Since the simulation results are all contained within the northeast quadrant of the plane then it was possible to calculate 95% uncertainty

intervals for cost-effectiveness and these uncertainty intervals are already included in the base case cost-effectiveness results reported in

Table 57. Indeed, because the joint distribution is relatively normally distributed, the

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

mean cost-effectiveness across the probabilistic results at [redacted] is almost identical to the deterministic result of [redacted]. The expected values across the simulated results are presented in Table 58 and show that the results of the PSA correspond closely with the deterministic results of

Table 57, indicating that the model is approximately linear. At a threshold value of

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

£20,000 per QALY the point estimate of the net health benefit (NHB) is [redacted] and the probability of being cost-effective is [redacted]. At a threshold value of £30,000 per QALY the NHB point estimate is [redacted] and the probability of being cost-effective rises to [redacted].

Table 58: Base case probabilistic cost-effectiveness results for the 12-month model

Technology	Cost	QALY
Increment*	[redacted]	0.024 (0.015 to 0.034)
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

*Adjusted for persistence
**95% uncertainty intervals from the probabilistic analysis
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio

Figure 19: Probabilistic results for the base case cost-effectiveness analysis presented on the cost-effectiveness plane

QALY=quality-adjusted life year

B.3.11.2 Deterministic sensitivity analysis

In this analysis, we have placed a strong focus on the statistical results of study 301 and 303. For this reason, a traditional one-way DSA of each individual parameter listed in Table 56 would not reveal very much as each ISI© parameter by itself contributes only a small part to the overall uncertainty. For this reason, we choose to present an atypical DSA that is in part based on the PSA. Where groups of parameters are estimated from regression equations then we allow all parameters in that group to vary while holding other parameters constant at their base case values. This allows us to construct the usual tornado diagram to summarise the influence of these parameter groups and this is presented in Figure 20 with the corresponding table showing the values added to Appendix N. As expected, it is the observed ISI© values of both study 301 and 303 that are the most influential set of parameters. The mapping parameters to EQ-5D are next most influential followed closed by the drop-out parameters.

Figure 20: Tornado diagram showing the influence of each (group of) parameter(s) on the incremental cost-effectiveness ratio of daridorexant

ISI=Insomnia severity index; IP=inpatient; GP=general practitioner; ER=emergency room; ICER=incremental cost-effectiveness ratio

B.3.11.3 Scenario analysis

In Figure 21 the scenarios reported in this section are reported as a forest plot and compared to the base case assumptions. PSA was undertaken for all scenarios.

Figure 21: Forest plot showing the Base Case results compared to other scenarios and subgroups

NHWS=Cerner Enviza National Health and Wellness Survey; SDS=Sheehan Disability Scale; ICER=incremental cost-effectiveness ratio

A detailed description of each scenario follows below including the tables of results. There are three important considerations for the scenario analyses.

1. The base case analysis is cautious and conservative. There were opportunities to make assumptions that would improve the cost-effectiveness, but we have preferred to take a more considered approach.

2. In particular, we have argued for a short-term model that focuses on the 12-months of the daridorexant clinical trial programme where we have the best information about its effectiveness. Taking a lifetime perspective that includes possible long-term mortality benefits improves the cost-effectiveness.

3. Subgroup analysis by moderate versus severe insomnia at screening is not statistically significant. Although there is a numerical advantage for the severe group, the additional uncertainty means that it is impossible to reliably distinguish the subgroups relative to the base case.

Best and worst-case scenario

As described in Figure 15 and Figure 18 above, alternative assumptions could have been made concerning the ISI© trajectories for no treatment. In the worst case, or most pessimistic scenario, full placebo adjustment from 301 through 303 would reduce the overall ISI© benefit (and corresponding QALY benefit) at the end of 12 months. However, in practice untreated patients would not receive the placebo effect observed in clinical trials. The most favourable, or optimistic scenario, therefore, is to compare daridorexant to a no treatment trajectory that continues at the original baseline.

The results for the worst-case are presented in Table 59 with the corresponding cost-effectiveness plane presented in Figure 22. For the worst case there is a [redacted] probability the intervention is cost-effective at the £20,000 threshold, rising to [redacted] at £30,000.

The best-case results are presented in Table 60 with the corresponding cost-effectiveness plane presented in *Adjusted for persistence

QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 23. For the best case the probability of being cost-effective is 100% at either threshold.

Table 59: Worst-case cost-effectiveness results for the 12-month model (placebo adjust 303)

Technology	Cost	QALY
No Treatment	£614	0.703
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.022
ICER	[redacted]	
Increment*	[redacted]	0.017
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 22: Probabilistic results for the worst-case cost-effectiveness analysis presented on the cost-effectiveness plane

QALY=quality-adjusted life year

Table 60: Best-case cost-effectiveness results for the 12-month model (no placebo adjustment)

Technology	Cost	QALY
No Treatment	£683	0.613
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.112
ICER	[redacted]	
Increment*	[redacted]	0.082
ICER*	[redacted]	
NHB (20k)*	[redacted]	

Technology	Cost	QALY
NHB (30k)*	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 23: Probabilistic results for the best-case cost-effectiveness analysis presented on the Cost-effectiveness plane

QALY=quality-adjusted life year

Inclusion of indirect costs

Although not included in the reference case analysis, indirect costs are an important consideration for insomnia treatment. The effects of insomnia can have severe impacts on productivity level and daily functioning, quantified in this analysis as absenteeism and presenteeism. As described in B.3.5.4 Miscellaneous unit costs and , two ways of generating productivity gains from treatment are explored: directly using SDS© data collected in study 301 and 303, and indirectly from the WPAI mapped to ISI© in the NHWS.

Directly estimated from the clinical trial programme (SDS©)

Starting with the directly observed SDS© measure of productivity, the addition of the productivity results from Table 55 to the base case cost-effectiveness results of

Technology	Cost	QALY
No Treatment	£624	0.691

Technology	Cost	QALY
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

Table 57 results in the cost-effectiveness results presented in Table 61. The associated probabilistic results are presented in QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 24 on the cost-effectiveness plane and these simulation results are used to calculate the confidence intervals in Table 63.

Table 61: Cost-effectiveness results for the 12-month model: including productivity losses estimated directly from SDS©

Technology	Cost	QALY
No Treatment	£2,071	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 24: Probabilistic results for the inclusion of directly observed SDS© productivity losses in the cost-effectiveness analysis presented on the Cost-effectiveness plane

QALY=quality-adjusted life year

The point estimate shows the treatment cost is almost entirely offset by the productivity benefits with just a small positive increment remaining. In the probabilistic analysis of
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Figure 24, uncertainty is such that 54% of the simulation results fall into the northeast quadrant. Commensurate with an upper limit of [redacted] per QALY, the probability of treatment being cost-effective at the £20,000 – £30,000 per QALY thresholds are [redacted] and [redacted], respectively.

Indirectly estimated by mapping WPAI to ISI© in the NHWS

For the indirect estimation of productivity losses via a mapping algorithm to ISI©, the results from Figure 4 for each ISI© point are applied to the base case results of

Technology	Cost	QALY
No Treatment	£624	0.691
Daridorexant	[redacted]	0.725

Technology	Cost	QALY
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER*	[redacted]	
NHB (20k)*	[redacted]	
NHB (30k)*	[redacted]	

Table 57 in order to generate the results presented in Table 62.

Table 62: Cost-effectiveness results for the 12-month model, including productivity losses estimated indirectly from WPAI to ISI© mapping in NHWS

Technology	Cost	QALY
No Treatment	£13,204	0.691
Daridorexant	[redacted]	0.725
Increment	[redacted]	0.034
ICER	[redacted]	
Increment*	[redacted]	0.024
ICER* (95% interval)	[redacted]	
NHB (20k)* (95% interval)	[redacted]	
NHB (30k)* (95% interval)	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Inclusion of productivity losses offset the costs of treatment leading to dominant situation (indicated by the negative ICER results and the highly positive NHB). This is further emphasised when looking at the simulation results on the cost-effectiveness plane (Figure 25) – [redacted] of the simulations ([redacted]) fall into the 'dominant' southeast quadrant with [redacted] falling into the northeast trade-off quadrant ([redacted]). Interestingly, the stronger (negative) correlation from the use of the mapping function is clearly apparent in the simulation results as shown Figure 25.

Figure 25: Probabilistic results for the inclusion of productivity losses in the cost-effectiveness analysis presented on the Cost-effectiveness plane

QALY=quality-adjusted life year

The direct and indirect approaches to productivity estimation led to different estimates of productivity losses. This may be due to a number of reasons. First, only employed people completed the WPAI, whereas all clinical trial participants were invited to complete the SDS©. The results from the WPAI are more reflective of the loss of productivity at work, while the SDS© results are generalisable to all types of activities (i.e., school or work or normal daily activities). Second, the WPAI asks respondents to report in number of hours and the SDS© in number of days. The former is thought to be more precise. In addition, the WPAI collects the number of hours worked, so that absenteeism and presenteeism are calibrated on individual work routine. Third, the WPAI assumes that the level of impairment due to the disease is the same every day. For the SDS©, the number of unproductive days is estimated and then multiplied by the level of underproductivity. Overall, the absolute estimates of WPAI and SDS© are likely not comparable but rather explaining different components of the loss of productivity: the WPAI focuses on work productivity with high precision and the SDS© measures overall loss of productivity.

It is noteworthy that the methodology of mapping to ISI© only uses the independent effect of ISI© on productivity as the estimated treatment effect such that despite the large differences in the absolute measures of productivity, the incremental productivity

gains associated with treatment are of the same general magnitude ([redacted] in productivity gains for SDS© when adjusted for persistence versus [redacted] for the comparable productivity gains from the WPAI mapping).

Lifetime cost-effectiveness model

As mentioned earlier, there are several reasons for presenting a short-term model as the base case of this submission. One of them was the lack of any claimed mortality benefit for daridorexant. Certainly, there is no evidence of a mortality effect in the short-term clinical trial programme. However, it is plausible that improved sleep could have long-term health benefits and there are some epidemiological associations that could form the basis for modelling (12). Since the long-term benefits of daridorexant are highly uncertain we chose to present a lifetime analysis only as a possible scenario. The lifetime model requires additional assumptions, many of which are highly uncertain, including:

- Relationship between improved sleep and long-term health outcomes
- That improved sleep through pharmacological treatment will achieve improved health outcomes as if they are naturally occurring
- Treatment persistence rates going far into the future.

The evidence of an association of long-term health outcome and sleep was reviewed in B.1.3.2 Epidemiology. Here we focus on one of the few epidemiological studies that estimated a relationship between duration of sleep and mortality risk. Yin and colleagues conducted a meta-analysis of studies and reported the relative risk of low sleep duration (<6hrs per night) and 6-7hrs sleep duration on mortality risk as 1.04 and 1.01 respectively compared to a reference sleep duration of 7hrs or above (12). This can be combined with the estimated improvement in sleep duration reported in study 301 (Table 14) to estimate the possible mortality benefits of daridorexant based on the average of 24 minutes increased sleep duration in study 301 (categorised to correspond to the definitions presented in Yin et al).

These epidemiological parameters and distributions across sleep duration categories are additional parameters for the lifetime model listed in Table 63. Other parameters required include the annual discontinuation rate and discount rates for both costs and

outcomes (QALYs). Not shown in Table 63, are the standard lifetables for the UK sourced from the Office of National Statistics (ONS), that form the basis of the background mortality rates in the lifetime model.

Table 63: List of additional parameters to extend the short term-model to a lifetime model with names, values, description and the distribution used for the probabilistic analysis

Name	Value	Distribution	Description
doRate	5%	beta	Annual Rate of dropout
btw67	1.01	lognormal	Relative risk of mortality for those getting less than 6hrs sleep
blw6hrs	1.04	lognormal	Relative risk of mortality for those getting 6-7hrs sleep
NT7plus	20%	Dirichlet	Proportion with sleep time of 7hrs plus without treatment
NT67	35%	Dirichlet	Proportion with sleep time 6-7hrs without treatment
NTblw6	46%	Dirichlet	Proportion with sleep time below 6hrs without treatment
D7plus	32%	Dirichlet	Proportion with sleep time of 7hrs plus on treatment
D67	33%	Dirichlet	Proportion with sleep time 6-7hrs on treatment
Dblw6	35%	Dirichlet	Proportion with sleep time below 6hrs on treatment
mnAge	50	NA	Average age at model entry
cDR	3.50%	NA	Annual discount rate for costs
oDR	3.50%	NA	Annual discount rate for QALYs

hr=hour; QALY=quality-adjusted life year; NA=not applicable

The lifetime model takes the short-term model presented as the base case analysis for the first year. In subsequent years, patients are assumed to experience the 'long-term' cost-effectiveness of daridorexant providing they persist with treatment. This is the cost-effectiveness that is apparent at the end of the first year for those remaining on treatment. While the average cost-effectiveness over the first year in the base case model was just below [redacted] per QALY, the cost-effectiveness for those completing one-year of treatment was [redacted]. It is this latter figure that is utilised for subsequent years for those remaining on treatment, but with an additional benefit obtained each year by avoiding mortality. The results of the lifetime model are presented in Table 64.

Table 64: Lifetime cost-effectiveness results

Technology	Cost	QALY
Increment*	[redacted]	0.327
ICER* (95% interval)	[redacted]	
NHB (20k)* (95% interval)	[redacted]	
NHB (30k)* (95% interval)	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

The point estimate of the ICER from the lifetime model did not differ substantially from the base case (Table 64) – this is because the lifetime model tends toward the 'long-term' results from the short-term 12-month model, while the inclusion of a mortality benefit has little impact on the point estimate. Furthermore, the meta-analysis informing the mortality benefits is rather weak evidence because it is not treated insomnia and shows a U-shape across sleep duration with longer than average sleep duration also predicting higher mortality (12). This is likely due to the direction of causation – with those illnesses associated with reduced mortality also associated with greater sleep duration. In addition, the lifetime model also results in greater uncertainty as illustrated by the cost-effectiveness plane (Figure 26).

Figure 26: Probabilistic results for the lifetime cost-effectiveness analysis presented on the Cost-effectiveness plane

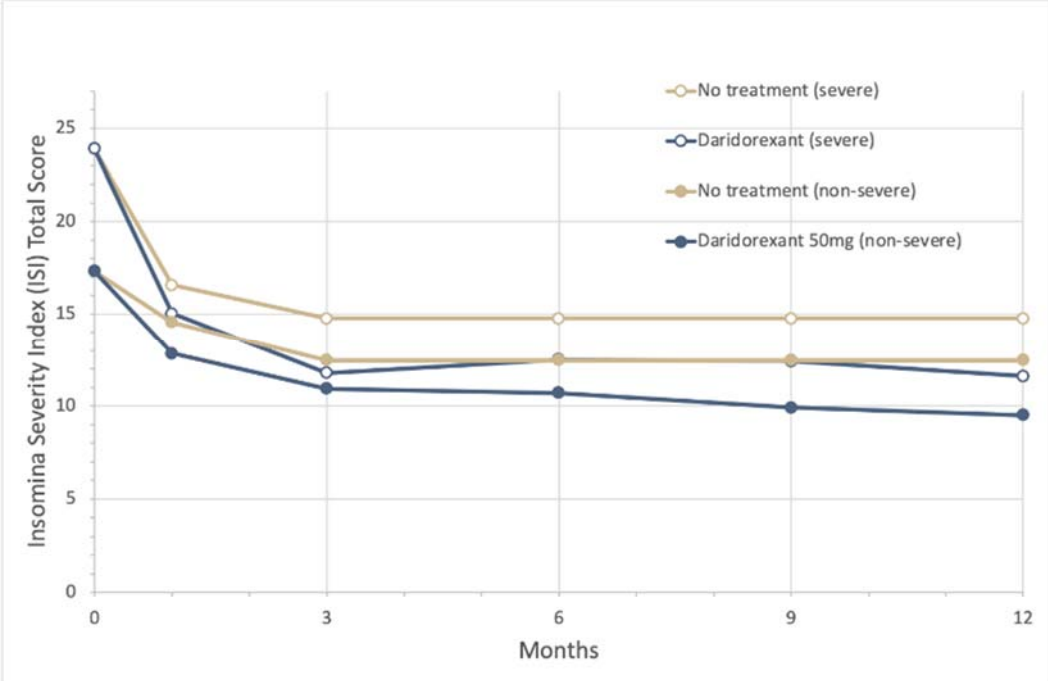
QALY=quality-adjusted life year

B.3.12 Subgroup analysis

Recognising the role of the ISI© in recruiting patients to the daridorexant clinical trial programme, we provide a subgroup analysis based on categorisation of the data into two groups: one with severe insomnia (ISI© >21) and the other with moderate insomnia at screening (ISI© 15–21). This categorisation was achieved by including main effects of treatment and subgroup, together with their interaction in the seemingly unrelated

regression analysis reported in B.2.9 Additional analysis of ISI©, Table 49. When combined with the stratification of study 303 by severity, the profile of ISI© modelled in the cost-effectiveness analysis is presented in Figure 27.

Figure 27: ISI© trajectories based on sub-groups of starting ISI©: severe (22-28) and moderate (15-21) at screening



ISI=Insomnia Severity Index

Although we report the subgroup analyses here, it is important to note that in neither of the regressions were the interaction terms in the regressions statistically significant (B.2.9 Additional analysis of ISI©, Table 49). Based on a traditional interpretation, the two groups would not be considered significantly different based on the observed data in study 301. Furthermore, the stratification of study 303 results in considerable uncertainty as can be seen from Figure 27 where the trajectories into the later period of the year that are informed by study 303 showed some volatility. In estimating these subgroup effects, we propagate the uncertainty reflected in all the parameters so that the insignificance of the interaction term and the volatility in study 303 stratification appear as uncertainty in the cost-effectiveness results.

The cost-effectiveness for the moderate and severe subgroups are presented in Table 65 and Table 66 respectively. The point estimate of the ICER favour the severe subgroup which has a higher baseline ISI© and a larger treatment effect. However,

the wide confidence intervals reflect the lack of significance of the interaction in study 301 and the volatility of the stratification in study 303.

Table 65: Cost-effectiveness results for the 12-month model: moderate (15-21) at screening subgroup

Technology	Cost	QALY
No Treatment	£607	0.711
Daridorexant	[redacted]	0.737
Increment	[redacted]	0.025
ICER	[redacted]	
Increment*	[redacted]	0.018
ICER* (95% interval)	[redacted]	
NHB (20k)* (95% interval)	[redacted]	
NHB (30k)* (95% interval)	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

Table 66: Cost-effectiveness results for the 12-month model: severe (22-28) at screening subgroup

Technology	Cost	QALY
No Treatment	£635	0.676
Daridorexant	[redacted]	0.709
Increment	[redacted]	0.033
ICER	[redacted]	
Increment*	[redacted]	0.025
ICER* (95% interval)	[redacted]	
NHB (20k)* (95% interval)	[redacted]	
NHB (30k)* (95% interval)	[redacted]	

*Adjusted for persistence
QALY=quality-adjusted life year; ICER=incremental cost-effectiveness ratio; NHB=net health benefit

The level of uncertainty is further illustrated on the cost-effectiveness plane (Figure 28). The extent of overlap between the groups is substantial despite the severe group being positioned in the slightly more cost-effective direction.

Figure 28: Severe (gold) and non-severe (blue) subgroup simulations on the cost-effectiveness plane

QALY=quality-adjusted life year

B.3.13 Benefits not captured in the QALY calculation

We acknowledge that EQ-5D is the preferred instrument for generating the QALYs required for cost-effectiveness analysis in the NICE reference case. In the absence of directly observed EQ-5D in the daridorexant clinical trial programme, we have presented a *de novo* mapping of ISI© to EQ-5D. This mapping has shown that insomnia disorder as measured by ISI© does correlate with EQ-5D and was used to estimate the QALYs presented in this submission.

Nevertheless, it is plausible that EQ-5D may not fully capture the impact of insomnia disorder on HRQoL. It has long been understood that EQ-5D may miss important dimension of HRQoL for some conditions – past research have explored the potential use of 'bolt-on' dimensions to capture missing dimensions. One of the most popular candidates for a bolt-on is fatigue, as fatigue is a feature of many health conditions including insomnia disorder (114). Perneger and Courvoisier examined possible missing dimensions from EQ-5D and identified separately fatigue/ energy and sleep as two dimensions that are poorly represented (115). Therefore, we believe it is reasonable to consider that the QALY estimates presented in this submission are an underestimate of the benefits of daridorexant on HRQoL.

B.3.14 Validation

Validation of the cost-effectiveness model occurred at several levels. Face validity was done through presentation of the underlying concept of the analysis with several clinical experts, health technology assessment experts and NICE through the decision problem meeting. This was especially important given that this submission focused on using a short-term model.

A formal advisory board brought together clinical experts and health economics experts to review the in-progress model (98).

Technical validation was undertaken by Avalon Health Economics. Following completion of the model the programming was tested and the results replicated by individuals not involved in the initial programming of the model.

B.3.15 Interpretation and conclusions of economic evidence

A de novo cost-effectiveness model for the economic evaluation of daridorexant for insomnia disorder was developed for this submission in close alignment with the NICE scope. Model inputs were primarily derived from confirmatory study 301 and extension study 303, including inputs for baseline characteristics, clinical outcomes, productivity gains and treatment discontinuation rates. Additional model inputs for unit costs and resource use were identified from NHS National Reference Costs. Health utilities were derived by mapping ISI© scores to EQ-5D based on an algorithm developed from NHWS – a nationally representative annual cross-sectional survey. The model was able to reproduce the results of study 301 and 303 over a period of 12 months. In absence of clinical trial evidence, further extrapolation was not undertaken for the base case analysis, but a lifetime model was presented in a scenario analysis.

In the base case analysis, over the 12-month time horizon, patients treated with daridorexant experienced improved insomnia disorder symptoms as reflected by the reduction in ISI© score leading to a higher QALY gain compared to those with no treatment (0.725 QALY vs. 0.691 QALY). Incremental costs were largely attributable to treatment acquisition cost of daridorexant, which was partially mitigated by the reduced HCRU. The base case analysis estimated a deterministic ICER of [redacted] per QALY and a probabilistic ICER of [redacted] per QALY, with both [redacted], suggesting that daridorexant offers a good use of NHS resources and should be

preferred over no treatment based on usual threshold values. The base case analysis presents an average for all patients initiating treatment over the first 12 months. However, it is important to note that, due to selective attrition, patients remaining on treatment at the end of 12 months enjoy a much better than average cost-effectiveness ratio of [redacted] per QALY representing the value of long-term treatment with daridorexant.

Scenario and sensitivity analyses showed that the cost-effectiveness model was most sensitive to the inclusion of productivity benefits, extent of placebo adjustment and ISI© scores. The uncertainty around ISI© score was expected, as this is the main clinical outcome representing treatment effect in the cost-effectiveness model.

One of the major impacts of insomnia disorder is daytime functioning impairment. Although excluded from the NICE reference case, current NICE guidance allows the presentation of productivity adjusted cost-effectiveness "if such costs may be a critical component of the value of the technology" (116). We argue that for insomnia disorder, the productivity benefits of treatment are a critical component of value because productivity is a major part of patients' daytime functioning. Productivity, as measured by the SDS©, was included in daridorexant's clinical trial programme. Furthermore, it was possible to map ISI© to WPAI using NHWS. Both methods of including productivity gave similar results and both methods showed that inclusion of productivity benefits of daridorexant completely offset the acquisition costs of treatment in the first year. For those remaining on treatment at 12 months, inclusion of productivity benefits shows cost-savings to society from daridorexant treatment.

As discussed in B.3.3.2 Base case plus best/ worst case scenarios of ISI© trajectory were based on the extent of placebo correction of ISI© scores. Although the base case argued for placebo adjustment for study 301 only, the reality is that withdrawal of treatment results in a return to baseline ISI© for all subjects – both active treatment and placebo. The true impact of daridorexant on insomnia disorder compared to no treatment is therefore represented by the best-case scenario comparing daridorexant to baseline ISI©. This resulted in an average 12-month cost-effectiveness of [redacted] per QALY falling to [redacted] per QALY gained for those remaining on treatment at 12-months.

While the reasons for a short-term model in the base case have been extensively discussed in B.3.2.1 Model structure, a lifetime horizon model was presented as a scenario analysis. There is some epidemiological evidence that normal range sleep duration has all-cause mortality benefit (e.g., reduction of cardiovascular stress, fewer falls, errors and accidents at work and in public places) but the impact is small and does not affect the ICER substantially. However, the lifetime model does improve the ICER towards the long-term cost-effectiveness of those remaining on treatment as the relatively higher first year ICER is offset by the improved long-term cost-effectiveness for those remaining on treatment.

In conclusion, the reference case cost-effectiveness analysis presented in this submission demonstrates that daridorexant represents a cost-effective use of NHS resources for the treatment of insomnia disorder in primary care. But this result is not revealing the true value to the NHS. Recognising that cost-effectiveness is improved for those remaining on treatment, that productivity losses are an important component of value for those with insomnia disorder and that lack of treatment means no placebo effect, means that the introduction of daridorexant as a treatment for insomnia disorder in England can result in substantial improvements to health and cost-savings at the societal level.

B.4 References

1. Idorsia Pharmaceuticals Ltd. Data on file. Idorsia insomnia market research: May 2022 GP Omnibus results. 2022.
2. Berger B, Kornberger R, Dingemanse J. Pharmacokinetic and pharmacodynamic interactions between daridorexant, a dual orexin receptor antagonist, and citalopram in healthy subjects. Eur Neuropsychopharmacol. 2021;51:90-104.
3. Summary of Product Characteristics. QUVIVIQ.
4. Treiber A, de Kanter R, Roch C, Gatfield J, Boss C, von Raumer M, et al. The Use of Physiology-Based Pharmacokinetic and Pharmacodynamic Modeling in the Discovery of the Dual Orexin Receptor Antagonist ACT-541468. J Pharmacol Exp Ther. 2017;362(3):489-503.
5. Janto K, Prichard JR, Pusalavidyasagar S. An Update on Dual Orexin Receptor Antagonists and Their Potential Role in Insomnia Therapeutics. J Clin Sleep Med. 2018;14(8):1399-408.
6. Quviviq. Daridorexant.: European Medicines Agency; [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/quviviq] Accessed on May 2022.
7. Eric Suni AR. Insomnia 2022 [Available from: https://www.sleepfoundation.org/insomnia] Accessed on May 2022.
8. Sateia MJ. International Classification of Sleep Disorders-Third Edition. CHEST. 2014;146(5):1387-94.
9. Substance Abuse and Mental Health Services Administration. (2016). Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health. CBHSQ Methodology Report. Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration Substance Abuse and Mental Health Services Administration, Rockville, MD.; 2016
10. Tolentino JC, Schmidt SL. DSM-5 Criteria and Depression Severity: Implications for Clinical Practice. Front Psychiatry. 2018;9:450.
11. Cormier R. Sleep Disturbances. In: Walker HK HW, Hurst JW, editor. Clinical Methods: The History, Physical, and Laboratory Examinations 3rd edition. Boston: Butterworths 1990.
12. Yin J, Jin X, Shan Z, Li S, Huang H, Li P, et al. Relationship of Sleep Duration With All-Cause Mortality and Cardiovascular Events: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. Journal of the American Heart Association. 2017;6(9):e005947.
13. Ishak WW, Bagot K, Thomas S, Magakian N, Bedwani D, Larson D, et al. Quality of life in patients suffering from insomnia. Innov Clin Neurosci. 2012;9(10):13-26.
14. Wilson S, Anderson K, Baldwin D, Dijk DJ, Espie A, Espie C, et al. British Association for Psychopharmacology consensus statement on evidence-based treatment of insomnia, parasomnias and circadian rhythm disorders: An update. J Psychopharmacol. 2019;33(8):923-47.
15. Substance Abuse and Mental Health Services Administration. Table 3.36, DSM-IV to DSM-5 Insomnia Disorder Comparison. CBHSQ Methodology Report. Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration Substance Abuse and Mental Health Services Administration, Rockville, MD.; 2016
16. ICD-11 for Mortality and Morbidity Statistics (Version : 02/2022). Insomnia disorders. 2022 [Available from: https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1038292737] Accessed on May 2022.
17. Andrew D. Krystal; Ruth M. Benca; and Thomas S. Kilduff. Understanding the Sleep-Wake Cycle: Sleep, Insomnia, and the Orexin System. J Clin Psychiatry. 2013;74(suppl 1):3-20.
18. Bollu PC, Kaur H. Sleep Medicine: Insomnia and Sleep. Mo Med. 2019;116(1):68-75.
19. Winkelman JW. Insomnia Disorder. N Engl J Med. 2015;373(15):1437-44.
20. Lane JM, Jones SE, Dashti HS, Wood AR, Aragam KG, van Hees VT, et al. Biological and clinical insights from genetics of insomnia symptoms. Nat Genet. 2019;51(3):387-93.
21. Hammerschlag AR, Stringer S, de Leeuw CA, Sniekers S, Taskesen E, Watanabe K, et al. Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. Nat Genet. 2017;49(11):1584-92.
22. Lind MJ, Gehrman PR. Genetic Pathways to Insomnia. Brain Sci. 2016;6(4).
23. Levenson JC, Kay DB, Buysse DJ. The pathophysiology of insomnia. Chest. 2015;147(4):1179-92.
24. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health. 2015;1(1):40-3.
25. Riemann D, Baglioni C, Bassetti C, Bjorvatn B, Dolenc Groselj L, Ellis JG, et al. European guideline for the diagnosis and treatment of insomnia. J Sleep Res. 2017;26(6):675-700.

26. Eric Suni NV. How Is Sleep Different For Men and Women? : Sleep Foundation; 2022 [Available from: https://www.sleepfoundation.org/how-sleep-works/how-is-sleep-different-for-men-and-women] Accessed on May 2022.
27. NICE. Depression in adults: recognition and management 2009 [Available from: https://www.nice.org.uk/guidance/cg90] Accessed on June 2022.
28. Luik AI, Machado PF, Siriwardena N, Espie CA. Screening for insomnia in primary care: using a two-item version of the Sleep Condition Indicator. Br J Gen Pract. 2019;69(679):79.
29. Buysse DJ, Thompson W, Scott J, Franzen PL, Germain A, Hall M, et al. Daytime symptoms in primary insomnia: a prospective analysis using ecological momentary assessment. Sleep Med. 2007;8(3):198-208.
30. Hays RD MR, Spritzer K, Drake C, Roth T. Psychometric properties of the Restorative Sleep Questionnaire and Daytime Consequences of Sleep Questionnaire. Associated Professional Sleep Societies (APSS-19th Annual Meeting); Denver (CO)2005.
31. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep. 1997;20(10):835-43.
32. Moul D, Pilkonis P, Miewald J, Carey T, Buysse D. Preliminary study of the test-retest reliability and concurrent validities of the Pittsburgh Insomnia Rating Scale (PIRS). Sleep. 2002;25:A246-A7.
33. Yeun EJ, Shin-Park KK. Verification of the profile of mood states-brief: cross-cultural analysis. J Clin Psychol. 2006;62(9):1173-80.
34. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2001;2(4):297-307.
35. Hudgens S, Phillips-Beyer A, Newton L, Seboek Kinter D, Benes H. Development and Validation of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ). Patient. 2021;14(2):249-68.
36. Kendzerska TB, Smith PM, Brignardello-Petersen R, Leung RS, Tomlinson GA. Evaluation of the measurement properties of the Epworth sleepiness scale: a systematic review. Sleep Med Rev. 2014;18(4):321-31.
37. Gagnon C, Bélanger L, Ivers H, Morin CM. Validation of the Insomnia Severity Index in primary care. J Am Board Fam Med. 2013;26(6):701-10.
38. Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. Health Econ. 2002;11(1):1-8.
39. Spronk I, Polinder S, Bonsel GJ, Janssen MF, Haagsma JA. Adding a fatigue item to the EQ-5D-5L improves its psychometric performance in the general population. J Patient-Rep Outcomes. 2022;6(1):1.
40. Idorsia Pharmaceuticals Ltd. Data on file. Mapping the ISI to the EQ-5D-3L QoL metric. Executive Technical Report. 10th June 2022.
41. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep. 2011;34(5):601-8.
42. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims: U.S FDA; 2009 [Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims] Accessed on May 2022.
43. Phillips-Beyer A, Kawata A, Kleinman L, Kinter DS. 0455 The Insomnia Daytime Symptoms and Impacts Questionnaire: An Analysis of Clinically Meaningful Change Using Phase 3 Clinical Trial Data. Sleep. 2022;45(Supplement_1):A201-A2.
44. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol. 1996;11 Suppl 3:89-95.
45. Sheehan KH, Sheehan DV. Assessing treatment effects in clinical trials with the discan metric of the Sheehan Disability Scale. Int Clin Psychopharmacol. 2008;23(2):70-83.
46. Mental health statistics: global and nationwide costs [Available from: https://www.mentalhealth.org.uk/statistics/mental-health-statistics-global-and-nationwide-costs] Accessed on May 2022.
47. Idorsia Pharmaceuticals Ltd. Data on File. BV_Insomnia-NHWS Results_UK_v1.0_23March2021. . Cerner Enviza2021.
48. Rosekind MR, Gregory KB. Insomnia risks and costs: health, safety, and quality of life. Am J Manag Care. 2010;16(8):617-26.
49. Li L, Gan Y, Zhou X, Jiang H, Zhao Y, Tian Q, et al. Insomnia and the risk of hypertension: A meta-analysis of prospective cohort studies. Sleep Med Rev. 2021;56:101403.

50. Hein M, Lanquart JP, Loas G, Hubain P, Linkowski P. Prevalence and risk factors of type 2 diabetes in insomnia sufferers: a study on 1311 individuals referred for sleep examinations. Sleep Med. 2018;46:37-45.
51. NICE Clinical Knowledge Summaries [Available from: https://cks.nice.org.uk/topics/insomnia/] Accessed on May 2022.
52. Morphy H, Dunn KM, Lewis M, Boardman HF, Croft PR. Epidemiology of insomnia: a longitudinal study in a UK population. Sleep. 2007;30(3):274-80.
53. Chung KF, Yeung WF, Ho FY, Yung KP, Yu YM, Kwok CW. Cross-cultural and comparative epidemiology of insomnia: the Diagnostic and statistical manual (DSM), International classification of diseases (ICD) and International classification of sleep disorders (ICSD). Sleep Med. 2015;16(4):477-82.
54. Uhlig BL, Sand T, Odegård SS, Hagen K. Prevalence and associated factors of DSM-V insomnia in Norway: the Nord-Trøndelag Health Study (HUNT 3). Sleep medicine. 2014;15(6):708-13.
55. Calem M, Bisla J, Begum A, Dewey M, Bebbington PE, Brugha T, et al. Increased prevalence of insomnia and changes in hypnotics use in England over 15 years: analysis of the 1993, 2000, and 2007 National Psychiatric Morbidity Surveys. Sleep. 2012;35(3):377-84.
56. Ohayon MM, Reynolds CF, 3rd. Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). Sleep Med. 2009;10(9):952-60.
57. Data on file. Real World Evidence of Burden of Illness for Insomnia in Europe (EUR5: France, Germany, United Kingdom, Italy and Spain), the United States, and Japan.
58. National Health and Wellness Survey (NHWS): Cerner Enviza; [Available from: https://www.cernerenviza.com/real-world-data/national-health-and-wellness-survey-nhws] Accessed on May 2022.
59. World Population Prospects 2019: United Nations; 2019 [Available from: https://population.un.org/wpp/] Accessed on May 2022.
60. Depression and Sleep: Understanding the Connection: Johns Hopkins Medicine; [Available from: https://www.hopkinsmedicine.org/health/wellness-and-prevention/depression-and-sleep-understanding-the-connection] Accessed on May 2022.
61. Coronavirus and depression in adults, Great Britain: January to March 2021 2021 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/coronavirusanddepressioninadultsgreatbritain/januarytomarch2021] Accessed on May 2022.
62. Data on file. UK Insights Deep Dive. 2022.
63. Stone KL, Ensrud KE, Ancoli-Israel S. Sleep, insomnia and falls in elderly patients. Sleep Med. 2008;9 Suppl 1:S18-22.
64. Reynolds CF, 3rd, O'Hara R. DSM-5 sleep-wake disorders classification: overview for use in clinical practice. Am J Psychiatry. 2013;170(10):1099-101.
65. Roth T. Insomnia: definition, prevalence, etiology, and consequences. J Clin Sleep Med. 2007;3(5 Suppl):S7-10.
66. Zhang Y, Ren R, Lei F, Zhou J, Zhang J, Wing YK, et al. Worldwide and regional prevalence rates of co-occurrence of insomnia and insomnia symptoms with obstructive sleep apnea: A systematic review and meta-analysis. Sleep Med Rev. 2019;45:1-17.
67. Advancing our health: prevention in the 2020s: HM Government; 2019 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/819766/advancing-our-health-prevention-in-the-2020s-accessible.pdf] Accessed on May 2022.
68. Hafner M, Stepanek M, Taylor J, Troxel WM, Van Stolk C. Why sleep matters - the economic costs of insufficient sleep: A cross-country comparative analysis. Santa Monica, CA: RAND Corporation; 2016.
69. Daley M, Morin CM, LeBlanc M, Grégoire JP, Savard J, Baillargeon L. Insomnia and its relationship to health-care utilization, work absenteeism, productivity and accidents. Sleep Med. 2009;10(4):427-38.
70. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011;21(10):718-79.
71. Kripke DF. Risks of Chronic Hypnotic Use. Madame Curie Bioscience Database [Internet] Austin (TX): Landes Bioscience; 2000-2013.
72. Heesch CB. The long-term use of sedative hypnotics in chronic insomnia. Mental Health Clinician. 2014;4(2):78-81.

73. Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults. NICE guideline [NG215]: NICE; 2022 [Available from: https://www.nice.org.uk/guidance/ng215] Accessed on May 2022.
74. Sleepio to treat insomnia and insomnia symptoms. Medical technologies guidance [MTG70]: NICE; 2022 [Available from: https://www.nice.org.uk/guidance/mtg70] Accessed on May 2022.
75. Summary of Product Characteristics. Circadin 2 mg prolonged-release tablets [Available from: https://www.ema.europa.eu/en/documents/product-information/circadin-epar-product-information_en.pdf] Accessed on June 2022.
76. Data on file. IQVIA Longitudinal patient data (LPD) 2021.
77. Data on file. EUCan HCP Databook - Country Breakdowns May 2022.
78. Taylor S AF, Burkinshaw P, Greaves F, Kelleher M, Knight, J PC, Tran A, White M, Marsden J. Dependence and withdrawal associated with some prescribed medicines. An evidence review. : Public Health England, London, 2019; 2020 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/940255/PHE_PMR_report_Dec2020.pdf] Accessed on May 2022.
79. Agravat A. 'Z'-hypnotics versus benzodiazepines for the treatment of insomnia. Progress in Neurology and Psychiatry. 2018;22(2):26-9.
80. Stewart SA. The effects of benzodiazepines on cognition. J Clin Psychiatry. 2005;66 Suppl 2:9-13.
81. Brett J, Murnion B. Management of benzodiazepine misuse and dependence. Aust Prescr. 2015;38(5):152-5.
82. Eric Suni AD. Melatonin and Sleep: Sleep Foundation; 2022 [Available from: https://www.sleepfoundation.org/melatonin#:~:text=Some%20people%20experience%20daytime%20sleepiness,the%20supervision%20of%20your%20doctor] Accessed on May 2022.
83. Hutka P, Krivosova M, Muchova Z, Tonhajzerova I, Hamrakova A, Mlyncekova Z, et al. Association of Sleep Architecture and Physiology with Depressive Disorder and Antidepressants Treatment. Int J Mol Sci. 2021;22(3).
84. Overview - Cognitive behavioural therapy (CBT): NHS; [Available from: https://www.nhs.uk/mental-health/talking-therapies-medicine-treatments/talking-therapies-and-counselling/cognitive-behavioural-therapy-cbt/overview/] Accessed on May 2022.
85. Matthews EE, Arnedt JT, McCarthy MS, Cuddihy LJ, Aloia MS. Adherence to cognitive behavioral therapy for insomnia: a systematic review. Sleep Med Rev. 2013;17(6):453-64.
86. Rob Newson AD. Cognitive Behavioral Therapy for Insomnia (CBT-I) 2022 [Available from: Cognitive Behavioral Therapy for Insomnia (CBT-I)] Accessed on June 2022.
87. Cunningham JEA, Jones SAH, Eskes GA, Rusak B. Acute Sleep Restriction Has Differential Effects on Components of Attention. Frontiers in Psychiatry. 2018;9.
88. Morgenthaler T, Kramer M, Alessi C, Friedman L, Boehlecke B, Brown T, et al. Practice parameters for the psychological and behavioral treatment of insomnia: an update. An american academy of sleep medicine report. Sleep. 2006;29(11):1415-9.
89. Dündar Y, Dodd S, Strobl J, Boland A, Dickson R, Walley T. Comparative efficacy of newer hypnotic drugs for the short-term management of insomnia: a systematic review and meta-analysis. Human Psychopharmacology: Clinical and Experimental. 2004;19(5):305-22.
90. Holbrook AM, Crowther R, Lotter A, Cheng C, King D. Meta-analysis of benzodiazepine use in the treatment of insomnia. Cmaj. 2000;162(2):225-33.
91. Glass J, Lanctôt KL, Herrmann N, Sproule BA, Busto UE. Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits. BMJ. 2005;331(7526):1169.
92. Chen PL, Lee WJ, Sun WZ, Oyang YJ, Fuh JL. Risk of Dementia in Patients with Insomnia and Long-term Use of Hypnotics: A Population-based Retrospective Cohort Study. PLOS ONE. 2012;7(11):e49113.
93. Louzada LL, Machado FV, Nóbrega OT, Camargos EF. Zopiclone to treat insomnia in older adults: A systematic review. Eur Neuropsychopharm. 2021;50:75-92.
94. Ferracioli-Oda E, Qawasmi A, Bloch MH. Meta-analysis: melatonin for the treatment of primary sleep disorders. PLoS One. 2013;8(5):e63773.
95. Wade AG, Ford I, Crawford G, McMahon AD, Nir T, Laudon M, et al. Efficacy of prolonged release melatonin in insomnia patients aged 55–80 years: quality of sleep and next-day alertness outcomes. Current Medical Research and Opinion. 2007;23(10):2597-605.
96. Mignot E, Mayleben D, Fietze I, Leger D, Zammit G, Bassetti CLA, et al. Safety and efficacy of daridorexant in patients with insomnia disorder: results from two multicentre, randomised, double-blind, placebo-controlled, phase 3 trials. Lancet Neurol. 2022;21(2):125-39.

97. Idorsia Pharmaceuticals Ltd. Data on file. Final CSR. ID-078A303: Multi-center, double-blind, parallel-group, randomized, placebo-controlled, three doses, 40-week extension to studies ID-078A301 and ID-078A302 to assess the long term safety and tolerability of ACT-541468 in adult and elderly subjects with insomnia disorder. 2021.
98. Idorsia Pharmaceuticals Ltd. Data on File. Ad Board 10Mar22 v2 FINAL
99. Idorsia Pharmaceuticals Ltd. Data on file. ID-078A301 CSR: Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder. 15 October 2020.
100. Yang M, Morin CM, Schaefer K, Wallenstein GV. Interpreting score differences in the Insomnia Severity Index: using health-related outcomes to define the minimally important difference. Curr Med Res Opin. 2009;25(10):2487-94.
101. Winkler A, Auer C, Doering BK, Rief W. Drug Treatment of Primary Insomnia: A Meta-Analysis of Polysomnographic Randomized Controlled Trials. CNS Drugs. 2014;28(9):799-816.
102. Carey TJ, Moul DE, Pilkonis P, Germain A, Buysse DJ. Focusing on the experience of insomnia. Behav Sleep Med. 2005;3(2):73-86.
103. Furukawa TA, Cipriani A, Leucht S, Atkinson LZ, Ogawa Y, Takeshima N, et al. Is placebo response in antidepressant trials rising or not? A reanalysis of datasets to conclude this long-lasting controversy. Evid Based Ment Health. 2018;21(1):1-3.
104. Enck P, Bingel U, Schedlowski M, Rief W. The placebo response in medicine: minimize, maximize or personalize? Nat Rev Drug Discov. 2013;12(3):191-204.
105. Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. Prevention & Treatment. 1998;1(2):No Pagination Specified-No Pagination Specified.
106. McCall WV, D'Agostino R, Jr., Dunn A. A meta-analysis of sleep changes associated with placebo in hypnotic clinical trials. Sleep Med. 2003;4(1):57-62.
107. Jiang B, He D, Guo Z, Gao Z. Dynamic features of placebo effects addressing persistent insomnia disorder: A meta-analysis of placebo-controlled randomized clinical trials. J Sleep Res. 2020;29(4):e12997.
108. Guidance on the use of zaleplon, zolpidem and zopiclone for the short-term management of insomnia [TA77]: NICE; 2004 [Available from: https://www.nice.org.uk/guidance/ta77] Accessed on May 2022 Accessed on
109. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep. 2011;34(5):601-8.
110. Dieperink KB, Elnegaard CM, Winther B, Lohman A, Zerlang I, Möller S, et al. Preliminary validation of the insomnia severity index in Danish outpatients with a medical condition. J Patient Rep Outcomes. 2020;4(1):18.
111. Jones, Karen C., Burns, Amanda (2021) Unit Costs of Health and Social Care 2021. Unit Costs of Health and Social Care . Personal Social Services Research Unit, Kent, UK, 185 pp. ISBN 978-1-911353-14-0.
112. National Cost Collection for the NHS [Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#nccdata2] Accessed on May 2022.
113. Office of National Statistics: Annual Survey of Hours and Earnings 2021 [Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/allemployeesashetable1] Accessed on June 2022.
114. Geraerds A, Bonsel GJ, Janssen MF, Finch AP, Polinder S, Haagsma JA. Methods Used to Identify, Test, and Assess Impact on Preferences of Bolt-Ons: A Systematic Review. Value Health. 2021;24(6):901-16.
115. Perneger TV, Courvoisier DS. Exploration of health dimensions to be included in multi-attribute health-utility assessment. International journal for quality in health care : journal of the International Society for Quality in Health Care. 2011;23 1:52-9.
116. NICE health technology evaluations: the manual Process and methods [PMG36]. Published: 31 January 2022 [Available from: https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation] Accessed on June 2022.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Daridorexant for treating insomnia disorder [ID3774]

Clarification questions

August 2022

File name	Version	Contains confidential information	Date
ID3774_Daridorexant_clarification questions_ACIC redacted.docx	1.0	No	2nd August 2022

Section A: Clarification on effectiveness data

Literature searches

A1. Please provide the full strategies used for the searches of ClinicalTrials.gov, World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) and all conference proceedings.

Response: Hand searches were conducted for ClinicalTrials.gov, World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) and all conference proceedings on the trial registry platform/conference websites using the keyword 'insomnia'. Additionally, the trial registry citations (WHO ICTRP and clinicaltrials.gov) were searched in the Cochrane CENTRAL database (Please see Table 4 of Appendix D for the Cochrane CENTRAL database search strategy; search conducted on 1st March 2022).

Table 1: Electronic data sources and their corresponding time limits for search

Data source	Website	Time limits
clinicaltrials.gov	https://clinicaltrials.gov/	None
WHO ICTRP	http://apps.who.int/trialsearch/Default.aspx	None
British Sleep Society	https://www.sleepsociety.org.uk/	Last two conferences
European Sleep Research Society, European Sleep Research Society	https://esrs.eu/	Last two conferences
Sleep and Breathing, Europe	https://sleepandbreathing.org/	Last two conferences (Could not be accessed)
Société Française de Recherche et Médecine du Sommeil (SFRMS)	https://www.sfrms-sommeil.org/	Last two conferences (Could not be accessed)
International Society for Pharmacoeconomics and Outcomes Research (ISPOR)	https://www.ispor.org/	Last two conferences

ISPOR Europe (EU)	https://www.ispor.org/	Last two conferences

A2. Please provide the date ranges for all databases searched (including start and end date for each resource used).

Response: Date ranges of all databases searched are provided below:

Table 2: Date ranges for all databases

Database	Search Date limit
Embase	No date limit: All hits up to Search
Medline	No date limit: All hits up to Search
Cochrane Library	No date limit: All hits up to Search
PsycInfo	No date limit: All hits up to Search

A3. Please confirm whether the Sleep and Breathing, Europe and the Société Française de Recherche et Médecine du Sommeil (SFRMS) conferences were searched. They are listed in the tables of included resources, however it is subsequently stated that *"Sleep and Breathing, Europe and the SFRMS required memberships for access and could therefore not be searched"*.

Response: We were not able to search the conference proceedings for Sleep and Breathing, Europe and the Société Française de Recherche et Médecine du Sommeil (SFRMS) due to lack of access.

A4. Please explain why conference proceedings were excluded from the Embase clinical effectiveness searches (Appendix D of the company submission (CS), Table 2). Although recent named conferences were searched, might a broader range of conferences with no date limit have provided additional useful information?

Response: The database search for Embase was designed broad to include any citations reporting on insomnia (without narrowing down to chronic

insomnia). Conference proceedings were excluded from the Embase clinical effectiveness searches due to a high volume of yield resulting from any conference proceedings reporting on 'insomnia', introducing a high number of irrelevant publications to screen. Hence, a targeted approach was followed by specifically hand searching conferences of interest in the past two years. It is standard practice to search for conference proceedings of preceding two years, as any study results published before would be reported in a peer review publication, which can be captured through database search. In instances where only one conference was conducted in the past two years, the search was extended to include last two meetings (E.g. European Sleep Research Society Conference conducted in 2018 and 2020).

A5. Please explain why search terms for cognitive behavioural therapy (CBT) were included in the cost-effectiveness searches (Appendices G-I), but not in the clinical effectiveness searches (Appendix D).

Response: For the clinical effectiveness SLR search, CBTi was not searched because the primary positioning of the drug is post-CBTI, and hence a comparison with CBTi would not be relevant in the line of management. This is in line with the British Association for Psychopharmacology (BAP) guidelines, "In case of treatment failure, unavailability of CBTi, or inability to engage with CBTi, pharmacological treatment with an evidence base should be offered"(1).

Further, it was known that there is no direct trial comparing daridorexant to CBTi. From previously published NMA/ITCs, there was an indication that there is huge heterogeneity amongst the RCTs of pharmaceutical interventions, making any form of indirect comparison difficult to interpret. Comparison of pharmaceutical intervention with non-pharmaceutical intervention would further add heterogeneity. Moreover, the delivery of CBTI intervention itself is likely to be highly variable between trials, and hence preclude the possibility of indirect comparison. Adding CBTi and related terms to the search strategy introduced a lot of noise in the search, a number of hits were not relevant to the research question. Restricting the search strategy to pharmaceutical interventions might have led to omission of a few RCTs on CBTi compared to other interventions

however, we believe none those would be able to establish comparative efficacy between CBTi and daridorexant directly or indirectly due to above mentioned heterogeneity issue.

The search terms for cost-effectiveness analysis included CBTi because it was anticipated that there would be very limited CEAs in this disease area hence, scope was not restricted to pharmaceutical interventions alone.

Decision problem

A6. Priority question. The population in the decision problem is insomnia disorder, but the indication for daridorexant is adult patients with insomnia characterised by symptoms present for at least 3 months and considerable impact on daytime functioning. Section B.1.3 of the CS states that *"chronic insomnia, also known as insomnia disorder, is defined as symptoms occurring for ≥3 nights per week for ≥3 months together with daytime impairment"*. The inclusion criteria for study 301 include an Insomnia Severity Index (ISI) score of ≥15.

a) Please confirm the precise definition of the population and provide a reference for that definition, if available.

b) Please define what those symptoms are and what is meant by daytime impairment in a way that such patients could be identified in clinical practice.

c) If the population in the decision problem is broader than that included in study 301 then please discuss the implications for the applicability

of study 301 to the National Institute for Health and Care Excellence (NICE) final scope as well as for clinical practice.

a) **Response:** The population specified in the decision problem is adults with insomnia disorder. This is based on the definition provided by Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®), which defines insomnia disorder as *"dissatisfaction with sleep quantity or quality, associated with difficulty* initiating *or maintaining sleep, or early morning awakening. Furthermore, the sleep disturbance is associated with significant social or functional distress or impairment. Sleep difficulty occurs at least 3 nights per week and is present for at least 3 months, and occurs despite adequate opportunity for sleep"(2).* Additionally, the DSM-5® criteria of insomnia disorder is largely consistent with the patient population indicated in the Summary of Product Characteristics (SmPC) for daridorexant, and the same DSM-5® criteria has been used to enrol patients in the pivotal trials of daridorexant (studies 301 and 302).

b) **Response:** The symptoms of chronic insomnia include problems of sleep initiation or maintenance despite adequate opportunities or circumstances of sleep which impacts daytime functioning (3).

For diagnosis of insomnia disorder, current diagnostic classifications, viz. DSM-5® and International Classification of Sleep Disorders, Third Edition (ICSD-3) not only include symptoms of sleep difficulties, but also complaints of significant distress, or daytime impairment (3). Since insomnia disorder is a subjective condition, its diagnosis solely depends on patients' experience of sleep difficulties and daytime impairment. The common symptoms of distress due to daytime consequences include somnolence, fatigue, daytime sleepiness, cognitive deficit, mood disturbance, reduced motivation, proneness for accidents, and impaired work or relationship functioning (4). These symptoms may serve as primary indicators of daytime functioning impairment in clinical practice.

Further, various patient-reported outcome instruments validated in clinical practice are available to assess patients' sleep habits and daytime

functioning impairment. This includes: Daytime Insomnia Symptom Scale (DISS), (5) the Daytime Consequences of Sleep Questionnaire (DCSQ), (6) the Functional Outcomes of Sleep Questionnaire (FOSQ), (7) the Pittsburgh Insomnia Rating Scale (PIRS), (8) the Profile of Mood States (POMS), (9) the Sleep Functional Impact Scale (SFIS), the Insomnia Severity Index (ISI) (10), the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) (11) and the Epworth Sleepiness Scale (ESS) (12). As noted in our submission clinical guidelines do not recommend the use of any specific PRO instrument to assess insomnia symptoms in clinical practice.

c) **Response:** The population in the decision problem (i.e., adults with insomnia disorder as per the DSM-5® criteria) is not expected to be broader than that of study 301. The use of ISI© score ≥15 as an inclusion criterion in study 301 is unlikely to impact the generalisability of the findings to the population in the decision problem since ISI© <15 represents subthreshold insomnia (13).

A7. Priority question. Table 2 of the CS states that some participants may be treated with 25 mg, but the CS seems to focus on those participants treated with 50 mg, e.g. the cost effectiveness analysis only includes the 50 mg dose. Table 2 also states that *"the treatment duration should be as short as possible. The appropriateness of continued treatment should be assessed within 3 months and periodically thereafter"*.

a) Please clarify that the decision problem populations should exclude patients for whom only the 25 mg dose is applicable. If not, please provide details on the criteria the dose was decided on, how many participants were treated with 25 mg as well as the baseline characteristics and results for participants treated with 25 mg and 50 mg, respectively.

b) Please clarify the expected treatment duration in clinical practice and whether this could be longer than the study 303 duration as well as the model time horizon of 12 months.

c) Please provide the stopping rules, i.e. precisely how it should be determined that a patients should stop treatment, either due to success or failure, and discuss the impact on relative effectiveness, if any.

a) **Response:** The decision problem population excluded patients treated with 25 mg once daily dosage of daridorexant as this dosage is only indicated for patients with moderate hepatic impairment or where there is co-administration of moderate CYP3A4 inhibitors.

b) **Response:** The currently recommended drug classes for insomnia disorder are indicated for only a short duration (<4 weeks for hypnotics, ≤13 weeks for melatonin). However, in clinical practice these drug classes are commonly used beyond their recommended duration. A UK insomnia market landscape analysis showed that, on average patients were on prescription drugs for [redacted] days in 2021 (14). Specifically, the average duration of therapy was [redacted] days for zopiclone, [redacted] days for melatonin and [redacted] days for amitriptyline (15). Given the chronicity of insomnia disorder, the treatment duration of daridorexant will likely be similar to or longer than these prescription drugs. Thus, the cost-effectiveness model estimates ICER for the full population over the first 12 months and those remaining on treatment after 12 months (lifetime scenario).

c) **Response:** With daridorexant, no formal stopping rules have been contrived. Per the SmPC the appropriateness of continued treatment should be assessed within 3 months of starting daridorexant and periodically thereafter. Primary care clinicians can monitor patient response and evaluate the need to continue treatment using established tools and approaches. Daridorexant's characteristic feature of quick onset and short

half-life allows treatment benefit to occur rapidly while on the medication; however, treatment effect stops when treatment stops, as demonstrated by the placebo run-out phase in-between study 301 and 303. Patients who remain on treatment are likely to accrue the greatest treatment benefit.

A8. Priority question. The comparator is stated to be *"established clinical management (including sleep hygiene advice) without daridorexant"* and the NICE clinical knowledge summary (CKS) notes to *"offer cognitive behavioural therapy for insomnia (CBT-I) as the first-line treatment for chronic insomnia in adults of any age"* where chronic insomnia is defined as insomnia of >3 months. However, the cost effectiveness analysis only has no treatment as comparator and CBT-I was not a comparator in the clinical effectiveness evidence (placebo is the comparator in the main analyses).

a) **Please elaborate why CBT-I was not included as a comparator.**

b) **If CBT-I is not a comparator, given that it is first line treatment then should the population in the decision problem be modified to 2nd line, i.e. after CBT-I?**

c) **If CBT-I is not a comparator, was it or should be used as concomitant therapy? If so, please compare the rate of use of CBT-I between the arms in study 301 between study 301 and clinical practice in the National Health Service (NHS) of England & Wales as well as discuss the implications of any discrepancy.**

d) **Please include CBT-I as a comparator in the clinical effectiveness and cost effectiveness analyses.**

a) **Response:** In study 301, CBTi was not a feasible comparator considering the study's randomised double-blinded design. This design was necessary to minimise the impact of confounders and effect modifiers when assessing the efficacy and safety of daridorexant. However, CBTi was allowed as a previous or concomitant therapy. As a concomitant therapy, CBTi was allowed only if it was initiated at least one month prior to Visit 3, wherein the subject agreed to continue CBTi throughout the study (16).

   In clinical practice and in line with available guidelines, CBTi is recommended and should be offered as a first-line treatment for patients with insomnia disorder. However, in cases where digital or face-to-face CBTi is inaccessible, or where a patient is unable to follow CBTi steps, or refuses CBTi, daridorexant may be considered as an alternative pharmacological treatment. Pharmacological therapy should be started after CBTi has been offered and therefore CBTi was not considered as a comparator of daridorexant. This was discussed in detail during scoping, with feedback from clinical experts and patient groups, resulting in the removal of CBTi as a comparator from the final scope. This was reconfirmed in the Decision Problem Meeting.

b) **Response:** According to the positioning of daridorexant specified in B.1.3.6, CBTi is the first-line treatment for insomnia disorder, and in these patients, daridorexant will serve as a second-line option if patients fail to respond to digital or face-to-face CBTi. CBTi should always be recommended as first-line treatment for insomnia disorder. However, considering issues with access or inability of patients to follow CBTi steps or if patients refuse CBTi, daridorexant may be administered as an alternative pharmacological treatment.

c) **Response:** In study 301, CBTi was allowed as a concomitant therapy. Only three randomised subjects (0.3%; 1 subject in each treatment group) were treated with CBTi at screening. Of the 927 subjects (99.7%) not using CBTi at screening, 25 subjects (2.7%; 11, 7, and 7 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported previous treatment failure with CBTi, 10 subjects (1.1%; 1, 5, and 4 subjects [daridorexant 25 mg, 50 mg,

and placebo, respectively]) reported no access/no therapist where subject lives, and 59 subjects (6.4%; 15, 24, and 20 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported no reimbursement for CBTi (16). This highlights that study 301 has insufficient data to support the use of daridorexant as a concomitant therapy to CBTi, since only 0.1% of subjects were on concomitant CBTi. This was reflected in the company's proposed positioning of daridorexant (Section B.1.3.6).

d) **Response:** CBTi was not specified as a comparator in the final scope of the decision problem, and this was discussed and agreed at the Decision Problem Meeting. Therefore, CBTi is not included as a comparator in the CS as per the positioning of daridorexant stated in A8 (a) and A8 (b).

A9. Priority question. In Table 1 of the CS, it is stated that *"while digital or face-to-face CBTi is recommended as the first-line treatment for insomnia disorder, it may not be suitable for or accessible to all patients. Daridorexant may thus be suitable for this group of patients as an alternative first-line treatment"*.

However, on reviewing the populations in studies 301 and 302 (a feeder trial of study 303), it is apparent that the populations have had minimal exposure to CBT, e.g. in study 301 0.3% of participants were receiving CBT at screening, 2.7% reported a previous failed CBT, and 87.9% of patients did not know CBT existed or were never offered CBT as a treatment option. The percentage of participants who had no access, interest or who refused CBT was 9.8% for all reasons combined.

a) Please comment on the appropriateness of using a largely CBT naïve population to justify the use of a pharmacological intervention as an

alternative to CBT when it is apparent that most participants have never had the opportunity to receive or reject CBT.

b) Is the population in the decision problem patients for whom CBT-I are not suitable or not accessible?

a) **Response:** While CBTi is the recommended first-line treatment for insomnia disorder, it is associated with certain limitations that bottlenecks its utilisation.

- Poor access and availability of face-to-face CBTi has been a longstanding problem (17).
- CBTi is resource intensive, and depending on the patient's need the number of sessions may vary from 6-8 (18).
- Adherence to CBTi is often poor as patients have to invest personal time and discipline to practise CBTi measures during and after the sessions (19).
- Inconsistent results arise from lack of standardised accredited training for resources administering CBTi (19).

These limitations lead to high refusal and failure rates with CBTi, which may be reflective of the population in study 301. In such cases, clinicians resort to alternative pharmacotherapies (benzodiazepines, Z-drugs, and melatonin) for immediate relief of insomnia symptoms. As described in the CS, hypnotics can effectively treat night-time symptoms of insomnia disorder such as sleep onset and/or sleep maintenance, but psychological dependence often leads to its prescription longer than their recommended duration as no long-term alternates exist in clinical practice (20). NICE's recommendation for Sleepio® (a digital self-help CBTi for the treatment of insomnia disorder) may significantly improve the limitations of access and cost with CBTi, but as highlighted by NICE there is limited clinical evidence to show the effectiveness of Sleepio® compared with face-to-face CBTi (17).

b) **Response:** In the decision problem, the population for daridorexant treatment includes patients for whom CBTi is inaccessible, unavailable or unsuitable i.e. as an alternative treatment. In addition, daridorexant may be used as second-line treatment, maintenance treatment, or for rapid symptom relief:

- For treatment-naïve patients who fail to respond to digital or face-to-face CBTi, daridorexant may be administered as a second-line treatment.
- For treatment experienced patients who have already completed standard of care including pharmacotherapy, daridorexant can be an alternative option.
- When longer-term management of insomnia symptoms (i.e., beyond 4 weeks) is required, daridorexant may be administered as maintenance treatment.
- When a patient is awaiting access to CBTi or a sleep specialist, daridorexant may be administered to provide rapid symptom relief.

A10. In the CS it is stated that [redacted] of patients refuse CBTi, or cannot access it, when recommended by their general practitioners (GPs). Among those who receive either face-to-face or digital CBT-I, [redacted] fail to achieve the desired results, leading to an overall CBT-I success rate of only [redacted].

Please provide the characteristics of these patients along with information to explain these values.

Response: The CBTi refusal and failure rates were obtained from a recent survey conducted among 1,002 GPs in the UK. Respondents were asked up to 12 questions regarding insomnia; this included the number of insomnia patients seen in the last 3 months, standard insomnia treatment algorithms for patients with insomnia disorder, availability and funding of CBTi, its refusal and failure proportions and referral to secondary care. No patient characteristics were

collected as part of the survey. Moreover, the NICE's assessment of Sleepio® highlighted the high dropout and failure rates with digital or face-to-face CBTi, which mentioned that the dropout rate was as high as 61.6% (21). This translates to a maximum success rate of 38.4%, which is close to the [redacted] reported in the GP survey presented in the CS.

A11. On page 25 of the CS, it is stated that *"multiple psychiatric and medical conditions are frequently associated with insomnia and may have a reciprocal relationship"*. It is further stated that *"approximately 50% of patients with insomnia also have mood (e.g., major depressive disorder) or anxiety disorders (e.g., PTSD)"*. The NICE final scope also states that *"insomnia is associated with comorbid conditions such as chronic obstructive pulmonary disease, heart failure, chronic pain, and psychiatric conditions (depression, anxiety, substance abuse, and post-traumatic stress disorder)"*.

However, studies 301 and 302 (a feeder into study 303) both exclude patients *"with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse…"* and study 303 excluded those with *"ECG findings"* meaning those with cardiac issues may have been excluded.

To what extent do these selection criteria restrict the generalisability of the trial populations to the chronic insomniac population at large, and in England and Wales specifically?

Response: The strict inclusion/exclusion criteria allowed the selection of a well-characterised insomnia population, in need of pharmacological intervention, thus being representative of insomnia disorder population. The company acknowledges that many patients in clinical practice are likely to have comorbidities, including neuropsychiatric disorders resulting in the use of various concomitant CNS-active medications; however, the need to exclude subjects with some comorbid conditions was driven by the importance of

limiting factors that could interfere with the optimal evaluation of the efficacy and safety of daridorexant. Since the underlying mechanisms of insomnia are thought to be the same in subjects with and without psychiatric disorders, including depression, the exclusion of these subjects does not affect the generalisability of the study results to insomnia disorder population at large, as well as to the population in England and Wales.

A12. The NICE final scope recommends that *"sleep hygiene advice"* should be attempted before continuing along the treatment pathway. Please provide details on the sleep hygiene measures that had been previously tried in the trial populations.

Response: The median time since insomnia diagnosis of all subjects in study 301 was 7.1 years. Therefore, it can be assumed that most subjects have attempted sleep hygiene advice prior to study enrolment. Information regarding sleep hygiene advice was not collected for the trial population of study 301, as it would be prone to recall bias given that sleep hygiene advice is usually attempted shortly after diagnosing insomnia disorder before continuing along the treatment pathway.

A13. Table 3 of the CS provides an *"assessment of commonly prescribed insomnia treatment in the UK based on characteristics of an ideal treatment for insomnia disorder"*.

Please provide a revised version of that Table in which daridorexant has been added.

Response: This table is now revised with daridorexant included and presented below.

Table 3: Assessment of commonly prescribed insomnia treatment in the UK based on characteristics of an ideal treatment for insomnia disorder

Characteristics (1, 3, 22)	Nitrazepam, Temazepam (23-26)	Zopiclone (23, 25-27)	Melatonin (28, 29)	Daridorexant (16, 30, 31)
Induces sleep rapidly	✓	✓	✓	✓
Maintains sleep throughout the night	✓	✓	x	✓
Preserves sleep architecture	x	x	✓	✓
Improves daytime functioning	x	x	x	✓
Indicated for long term use	x	x	x	✓
No next-day residual effects	x	x	x	✓
No risk of dependence/ tolerance	x	x	✓	✓
No rebound insomnia / withdrawal upon discontinuation	x	x	✓	✓
Appropriate for adults and elderly	x	x	x	✓
Minimal important interactions	x	x	✓	✓

UK = United Kingdom.
Note: Fulfilment of each characteristic is based on non-comparative evidence; therefore a direct comparison should not be made between the drugs.

Systematic literature review (SLR)

A14. Priority question. Study 303 appears to be absent from the SLR, e.g. Table 8 of Appendix D of the CS, even though it is in the CS. Please explain this omission.

Response: Study 303 did not meet the SLR requirements due to the study design issues. In this extension study, subjects who had completed the study treatment and run-out period for studies 301 and 302 were re-randomised to receive either placebo or 25mg daridorexant in a 1:1 ratio. Including re-randomised patients would bias the results due to double counting same patients hence, this study was excluded from the SLR.

A15. Priority question. The PICO (population, intervention, comparator(s), outcome(s)) criteria used in the identification of evidence lists "placebo" and "other active agent" as comparators. However, the NICE final scope defines the appropriate comparators as *"established clinical management (including sleep hygiene advice) without daridorexant"* thus the criteria used in the identification of evidence would exclude relevant studies with appropriate and current non-pharmacological clinical therapies (this question is linked to question A8).

Please provide justification for this choice of comparator.

Response: As per our response to question A8, daridorexant is a pharmacological treatment positioned in second-line after interventions such as sleep hygiene and CBTi. Hence, the comparators of interest for this SLR were placebo or active agent.

A16. In Appendix D, it is stated that *"one researcher extracted data from the included papers into the DET, which was then validated by a second, senior investigator".*

Please provide further information on how disagreements and discrepancies in the data extraction process were resolved.

Response: Once the extractions were validated, these were sent back to the researcher who had performed the original extractions to make required changes. Any disagreements between the extractor and validator were brought forward and were resolved by a third, more senior investigator who reviewed the disagreement and provided a final decision.

A17. On page 15 of Appendix D, it is stated that *"Evidera conducted a high-level comparability assessment"*. This process appears to have been conducted to identify and prioritise trials for potentially conducting a network meta-

analysis (NMA). It further states that *"these trial characteristics did not lead to exclusion from the SLR"* but then later on it states "*upon determination of the trials that were most likely to be suitable for NMA, a subset of trials underwent full extraction".*

a) Please confirm that no trials were excluded from the results of the Evidera comparability assessment.

b) Please verify that all trials, which met the PICO criteria and were included at full screening stage, did undergo data extraction.

c) Please provide full details on the methods and processes that were utilised to include/exclude trials.

a) **Response:** We confirm that no trials were excluded from the results of the Evidera comparability assessment.

b) **Response:** A top-level extraction was performed for all included studies. This top-level extraction – 2 step extraction – provided sufficient information and data to determine comparability assessment - information on trial, patient, treatment characteristics, and outcome availability (i.e., tagging for the presence of relevant outcomes) were recorded. Studies (as listed in Table 7, Appendix D) were then de-prioritised if they did not include a treatment arm using a licensed dose of a treatment of interest, if they did not report comparable data, or they did not report data in populations of most interest.

c) **Response:** Search results were uploaded to Distiller Systematic Review (DSR) software, an internet-based program that facilitates collaboration among reviewers during the study selection process. Screening followed a two-stage process:

Level 1: Titles and abstracts of studies identified by the search strategies were reviewed independently by two researchers to determine eligibility

according to the inclusion and exclusion criteria. Disagreements between the reviewers were resolved by a third reviewer, as needed.

Level 2: Articles deemed eligible during level 1 screening were reviewed independently by two researchers as full texts to determine eligibility according to the selection criteria.

As reported in section D.1.2.2 of Appendix D, disagreements between the reviewers were resolved by a third reviewer, as needed. After identifying the articles recommended for inclusion (in accordance with the PICOS criteria), a list of accepted studies and articles excluded at the full-text screening level were gathered, organized by reasons for exclusion, and the flow of studies was documented in a PRISMA diagram

Clinical effectiveness evidence

A18. Priority question. Please provide full details of the anticipated marketing authorisation for daridorexant.

Response: Currently, marketing authorisation approval of daridorexant is still pending for MHRA. In March 2021, marketing authorisation application for daridorexant was submitted to the EMA. A positive CHMP opinion was issued in February 2022, and marketing authorisation was approved on 29th April 2022 by EMA for *"the treatment of adult patients with insomnia characterised by symptoms present for at least 3 months and considerable impact on daytime functioning (32)."* The marketing authorisation by MHRA is anticipated to be consistent with that of EMA.

A19. Priority question. NCT02839200 (Dauvilliers et al. 2020) is included in the SLR, e.g. Table 8 of Appendix D of the CS, but not in the CS, even though it appears to be relevant, as it compares 50 mg daridorexant to placebo.

Please explain why this reference is not included in the CS and provide an addendum to the CS, if appropriate.

Response: As elaborated in Table 4 of CS, the clinical trial programme of daridorexant included two phase II studies, one of which was NCT02839200. The study was a 6-arm randomised trial, that included 4 dosages of daridorexant, zolpidem and placebo. Primarily, this trial assessed dose response relationship between 5, 10, 25 and 50 mg dose of daridorexant and thus, was not designed to evaluate efficacy and safety of daridorexant compared with placebo due to the small sample size utilised in the study. Therefore, this study was not found to be relevant for the appraisal.

A20. Priority question. A large number of outcomes are presented in the clinical evidence. Each outcome in the NICE final scope is therefore populated by several outcomes that evaluate similar parameters. This increases the probability of observing significant findings.

a) Please provide an overview with definitions of all outcomes used in the CS and explain how these relate to the outcomes listed in the NICE final scope.

b) Please provide a prioritisation of the outcomes within each category of NICE final scope outcome, with a clear rationale.

a) **Response:** The table below defines all outcomes used in the CS and their relationship with those listed in the NICE final scope.

Table 4: Definitions of outcomes used in CS and NICE final scope

Outcomes used in the CS	Definitions of outcomes used in the CS	Outcomes listed in NICE final scope
Improvement of night-time symptoms of insomnia	WASO (sleep maintenance), LPS (sleep onset), subjective TST (sleep time)	Resolution of symptoms
Changes in sleep architecture and sleep efficiency	Time to fall asleep, number of awakenings during the night and duration of TST by sleep stage/quarter of the night, depth of sleep	Changes in sleep patterns and architecture
Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function	Quality of sleep, daytime alertness and ability to function as assessed by visual analogue scale	Sleep quality
Daytime functioning as measured by IDSIQ total	Daytime impact of insomnia on three dimensions:	Daytime alertness

score, sleepiness, alert/cognition and mood domain score	physical (sleepiness domain), cognitive (alert/cognition domain), and affective (mood domain)	
	Recurrence of insomnia was not assessed	Recurrence of insomnia
Adverse effects of treatment (next-day residual treatment effects and memory impairment)	Withdrawal symptoms, Rebound insomnia, Next-morning residual effect and daytime sleepiness	Adverse effects of treatment (including residual daytime sedation and memory impairment)
Indirectly by mapping ISI© to EQ-5D	No specific questionnaire for HRQoL. Combination of the patient-reported assessments of sleep quality (using a Visual Analogue Scale), daytime functioning (the IDSIQ questionnaire), and insomnia severity (the ISI questionnaire)	HRQoL

b) **Response:** The ISI© should be prioritised among all the outcomes presented in the CS as it is the key effectiveness parameter of the economic model. Given the complexity of assessing treatment outcomes in insomnia disorder, it is challenging to prioritise all other outcomes within each category of the NICE final scope since all outcomes within a category should be considered in totality and therefore carry equal importance when evaluating the clinical benefit of daridorexant. This is supported by a number needed to treat (NNT) analysis of the key endpoints of study 301 (i.e., WASO, LPS, sTST, IDSIQ and ISI©). The results of the NNT analysis show that all key endpoints have comparable NNTs at month 3, as indicated by the overlapping confidence intervals (Table 5) (33).

Table 5: Number needed to treat values for the responder analysis based on sTST, LPS, WASO, ISI and IDSIQ at Month 1 and Month 3 for daridorexant compared with placebo (33)

Variable	Threshold response definition	1 month	3 months
		Daridorexant 50mg NNT, mean (95% CI)	Daridorexant 50mg NNT, mean (95% CI)
sTST	Change from baseline of ≥ 55 min	[redacted]	[redacted]
LPS	LPS < 20 min at	[redacted]	[redacted]
WASO	WASO < 30 min at	[redacted]	[redacted]
ISI	Change from baseline of ≤ -7 points	[redacted]	[redacted]
	Total score ≤ 7 points	[redacted]	[redacted]
IDSIQ	Change from baseline in sleepiness domain score of ≤ -8 points	[redacted]	[redacted]
	Change from baseline in sleepiness domain score of ≤ -4 points	[redacted]	[redacted]
	Change from baseline in alert/cognition domain score of ≤ -12 points	[redacted]	[redacted]
	Change from baseline in alert/cognition domain score of ≤ -9 points	[redacted]	[redacted]
	Change from baseline in mood domain score of ≤ -7 points	[redacted]	[redacted]
	Change from baseline in mood domain score of ≤ -4 points	[redacted]	[redacted]
	Change from baseline in total score of ≤ -25 points	[redacted]	[redacted]
	Change from baseline in total score of ≤ -17 points	[redacted]	[redacted]

A21. No outcome data appear to be provided for the outcomes of recurrence of insomnia (NICE final scope), rebound insomnia (company list of outcomes), health related quality of life (except IDSIQ), quality of sleep (company list of outcomes), depth of sleep (company list of outcomes), daytime alertness (company list of outcomes) and daily ability to function (company list of outcomes).

If these are present in the current results in the CS or Appendix M, please identify the exact locations of these data, or add these data if necessary.

Response: Results for all the listed outcomes in the NICE scope and the company list of outcomes are provided in the CS, Appendix F or Appendix M. The table below indicates the exact locations of the data in question.

Please note:

- Recurrence of insomnia (NICE final scope) was not directly assessed in the trial subjects who experienced a treatment effect but those who subsequently discontinued treatment.

- In the clinical trials presented in the CS, HRQoL was assessed via IDSIQ, and no other instruments were utilised. HRQoL for the CEM was derived indirectly using ISI® scores collected from the trials mapped to EQ-5D, as described.

Table 6: Company list of outcomes and their corresponding location in the CS/Appendix M/Appendix F

Company list of outcomes	Page number and Table number of results (as reported in CS/Appendix M/Appendix F)
Rebound insomnia	Study 301: Appendix F, Section F.1.1.4, Table 6 Study 303: Appendix F, Section F.1.2.4, Table 12
Quality of sleep	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Depth of sleep	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Daytime alertness	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Daily ability to function	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3

A22. Neither study included participants from the UK, as Canada, Denmark, Germany, Poland, Spain, Switzerland, and the United States of America enrolled and randomised participants. The baseline characteristics tables

showed that participants in study 301 were approximately 1% Asian, 9.5% Black and 89.5% White, and that participants in study 303 were approximately 1% Asian, 8.5% Black and 89.5% White. This is different to the UK population as measured in the 2011 census [Population of England and Wales - GOV.UK Ethnicity facts and figures (ethnicity-facts-figures.service.gov.uk)], where 7.5% of the population are Asian, 3.3% of the population are Black and 86% of the population are White (the analogous information from the 2021 census is not currently available). This population difference could potentially therefore have an impact on applicability. There is no evidence from the sub-group analyses for study 303 that ethnicity is an outcome modifier, but ethnicity was only evaluated as a sub-grouping variable for subjective total sleep time (sTST) and IDSIQ in that study and was not evaluated as a sub-grouping variable for any outcome in study 301.

a) In the light of this, please comment on the generalisability of the trial population characteristics to the patient population in England and Wales.

b) Please provide data sub-grouped for ethnicity for all primary and secondary outcomes in both study 301 and 303.

a) **Response:** The company acknowledges the difference in ethnic distribution between the trial population and the patient population in England and Wales. However, as highlighted by the EAG, there is no evidence from the subgroup analyses for study 303 that ethnicity is an outcome modifier. Although this was only evaluated for sTST and IDSIQ, the company expects this to be applicable to all other primary and secondary endpoints.

b) **Response:** The small sample size of the Asian and Black subgroups precluded meaningful comparison of all primary and secondary outcomes across ethnic subgroups. As mentioned in the response to A22 (a), based

on the subgroup analysis for study 303, there is no evidence that ethnicity is an outcome modifier for sTST and IDSIQ and the company expects this to be applicable to all other primary and secondary endpoints.

A23. Given that insomnia disorder is associated with various comorbid conditions such as chronic obstructive pulmonary disease, heart failure, chronic pain, and psychiatric conditions (depression, anxiety, substance abuse, and post-traumatic stress disorder), please provide details on the clinical characteristics of any other pathologies present in the trial populations.

Response: In study 301, previous psychiatric disorders were reported for 54 subjects (5.8%), of which the most common was depression (24 subjects, 2.6%); additionally, major depression was reported for 7 subjects (0.8%) and anxiety for 4 subjects (0.4%). Previous nervous system disorders were reported for 29 subjects (3.1%), of which the most common was migraine (6 subjects, 0.6%).

Study concomitant medical conditions (excluding conditions and symptoms related to insomnia) were reported for 646 subjects (69.5%) and were balanced across the treatment groups. Table 7 illustrates the study concomitant medical conditions by primary system organ class and preferred term in the overall population of study 301 (16).

Table 7: Study concomitant medical conditions by primary system organ class and preferred term (16)

System Organ Class **Preferred Term**	**Total N=930** **n (%)**
Psychiatric disorders	43 (4.6%)
Tobacco abuse	15 (1.6%)
Anxiety	8 (0.9%)
Depression	4 (0.4%)
Nervous system disorders	121 (13.0%)
Headache	52 (5.6%)
Migraine	22 (2.4%)
Somnolence	12 (1.3%)

Metabolism and nutrition disorders	234 (25.2%)
Hypercholesterolaemia	74 (8.0%)
Obesity	70 (7.5%)
Type 2 diabetes mellitus	43 (4.6%)
Vascular disorders	223 (24.0%)
Hypertension	207 (22.3%)
musculoskeletal and connective tissue disorders	198 (21.3%)
osteoarthritis	73 (7.8%)
back pain	38 (4.1%)
Endocrine disorders	87 (9.4%)
hypothyroidism	72 (7.7%)

A24. No comparison was made across study arms for the number of participants using allowed treatment options, e.g. CBT-I. These had potential to be confounders if they differed between arms.

Please provide data on the numbers using allowed treatment options, e.g. CBT-I.

Response: CBT for any indication was only allowed if the treatment started at least one month prior to Visit 3 and the subject agreed to continue CBT throughout the study. In study 301, only three randomised subjects (0.3%; 1 subject in each treatment group) were treated with CBTi during the study. Thus, CBTi was not expected to be a confounder in the analyses (16).

Other therapies considered necessary for a subject's well-being was allowed during the study 301; however, the use of these therapies at baseline (study 301) and at start of double-blind treatment (study 303) was balanced across the treatment groups and were not expected to contribute as confounding factors in the efficacy and safety analyses (16).

The study design of daridorexant clinical trial excluded patients with acute or critical pathologies to prohibit the use of non-sedating antihistamines, opioids/narcotics, centrally acting muscle relaxants with psychotropic effects, pseudoephedrine, and inhaled or nasal corticosteroids. Further, randomization

of trial population ensured demographic and clinical characteristics of patients balanced confounding factors across the treatment arms (16).

A26. It is unclear why some validated and commonly used measurement tools were not used e.g. the Pittsburgh Sleep Quality Index.

a) Why was sleep duration was not objectively measured with e.g. wearable devices/actigraphy?

b) Why was total sleep time was only measured subjectively?

c) Why was sleep onset latency was not measured?

a) **Response:** The BAP guidelines recommend comprehensive assessment of subjective symptoms of insomnia disorder; objective measures, such as wearable devices/actigraphy or polysomnography (PSG) are indicated if sleep disorders such as sleep apnoea or narcolepsy are suspected (1). While wearables/actigraphy makes it convenient for trial subjects to measure sleep measures objectively, it tends to be less accurate than PSG and may not be sufficiently sensitive to detect changes in sleep parameters over time. In addition, as actigraphy assesses sleep based on movement, it is less accurate when evaluating fragmented sleep, reduced sleep time and/or restless sleep commonly seen in patients with insomnia disorder (34, 35).

b) **Response:** Total sleep time (TST) was assessed both subjectively and objectively. But the objective assessment of TST was an exploratory efficacy endpoint of study 301 (16). TST was defined as the time scored as non-awake from lights off to lights on, as determined by PSG. The results of the objective measure are presented in Section B.2.4.4, Table 15.

c) **Response:** Sleep onset latency was measured as objective subjective endpoints in study 301. Objectively it was assessed as a primary endpoint, latency to persistent sleep (LPS) and subjectively as an exploratory

endpoint, latency to sleep onset (LSO) in studies 301 (16) (Section B.2.4.1 and B.2.4.4).

LPS was the time from start of recording to the beginning of the first continuous 20 epochs (i.e., 10 min) scored as non-awake. Subjective LSO was the time reported by the subject in answer to the sleep diary questionnaire "How long did it take you to fall asleep?" (16)

A27. As per table 10 of the CS, *"900 subjects provided at least 90% power to detect an effect size of 0.37 for testing nine independent null hypotheses"* – however 0.37 corresponds to a small to medium effect size.

Please justify any clear clinical benefit for patients presenting with insomnia, e.g. by providing relevant references.

Response: There is a lack of evidence to support the use of a particular measure or combination of measures to demonstrate a clear clinical benefit for patients presenting with insomnia. Instead, an extensive list of outcomes was presented in the CS to provide a holistic assessment of the efficacy and safety of daridorexant. The company has established a meaningful threshold of 55 minutes for sTST compared to baseline using the dose response curve from a phase 2 study (36). It is challenging to establish a meaningful threshold compared to placebo since placebo effects are often large in insomnia studies.

A28. Please define the intention-to-treat analyses used, and for which outcomes.

Response: Intention-to-treat population was defined as all participants who were randomly assigned to a double-blind study treatment. In order to adhere to the intention-to-treat principle as much as possible:

- Subjects were evaluated according to the treatment and strata they were assigned to, which may differ from the treatment they received;
- All available data were included.

Intention-to-treat population was analysed in study 301 for the primary and secondary endpoints which included, objective assessments of WASO and LPS, and subjective assessments of TST and IDSIQ sleepiness domain.

A29. Page 20 of the CS lists Insomnia Severity Index© as a tool to measure the global assessment of insomnia severity. However, page 111 of the CS states that that same tool was used to quantify health related quality of life (HRQoL) in both studies 301 and 303. Please provide HRQoL results using validated tools such as EuroQol-5D (EQ-5D).

Response: In studies 301 and 303, HRQoL was not assessed directly with HRQoL instruments such as EQ-5D; instead the ISI© was used to assess and monitor insomnia severity at baseline and at various timepoints after administration of study treatment. The Cerner Enviza NHWS was utilised to develop a mapping algorithm (Section B.2.9, CS). As EQ-5D was not included in the clinical study, utility was captured indirectly through mapping from ISI© using the mapping algorithm.

A30. As per the final scope by NICE, *"availability and cost of biosimilar and generic products should be considered"*.

Please provide a rationale for not including/performing subgroup analyses on these variables.

Response: Pharmacotherapy is not recommended for long-term management of insomnia disorder. Most of the recommended short-term drugs for insomnia disorder are available as generic products. These are not considered as comparators of daridorexant, per the scoping and DPM discussions. Consequently, these analyses were not included in the CS.

A31. Please confirm that the latest data cut-off was 22 July 2020 and provide newer data, if available, in an addendum.

Response: The company confirms that the latest data cut-off was 22 July 2020 (31).

Section B: Clarification on cost-effectiveness data

Model structure

B1. Priority question. The model type is not explicitly stated (e.g. decision tree, state-transition model) nor is a detailed description provided regarding the model implementation and the related assumptions.

a) Please explicitly state and justify the model type that was used.

b) Please provide a detailed description regarding the model implementation and the related assumptions regarding:

1) Interpolation and combination of costs and effects estimated for the different time points, i.e. months listed in 'Model | 1 Year' row 11

2) Calculation and implementation of the ISI score, separately per treatment, including the difference between ISI and mnISI

3) Calculation and implementation of the utility score, separately per treatment, including the difference between EQ-5D and mnEQ-5D

4) Calculation and implementation of GP, emergency room and inpatient resource use, separately per treatment

5) Calculation of number of daridorexant 50 mg administrations, average per specified time unit

6) **Calculation and implementation of drop out adjustment (in 'Model | 1 Year'!C61:J71)**

7) **Justification for the drop out adjustment (in 'Model | 1 Year'!C61:J71)**

8) **Calculation of the total quality-adjusted life years (QALYs), separately per treatment, including a justification for not using the values in the following cells 'Model | 1 Year'!D48:H49 for estimating the incremental cost effectiveness ratio (ICER)**

9) **Calculation of the total costs, separately per treatment**

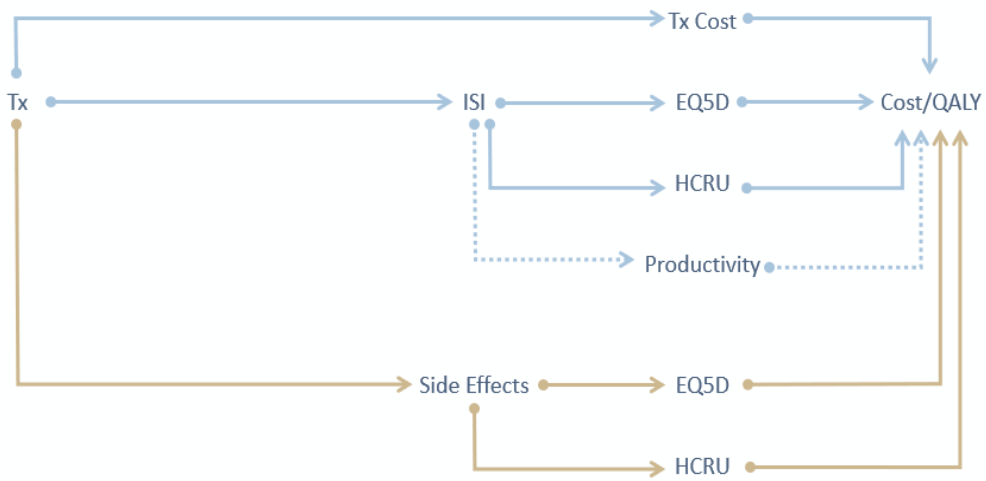
10) **Calculation of the ICER (in 'Model | 1 Year' rows 74:87)**

c) **Please provide a detailed description of specific model characteristics, e.g. half cycle correction, cycle time.**

d) **Please provide a similar level of detail (including justifications) regarding the estimation of input parameters (CS Table 63), model type, model implementation, model assumptions and model characteristics for the *"Lifetime cost-effectiveness model"*. This should include detailed explanations of the calculations and assumptions in the "Model | Lifetime" worksheet columns E:Q.**

a) **Response:** This model does not confirm to the standard 'decision analytic' type models mentioned. It is not a tree, nor does it have 'states'. In section B.3.2.1 where the model structure is described we call it a 'Pathway' model and present Figure 1 (reproduced below for your convenience) to illustrate the pathways.

Figure 1: Pathway from treatment to value via ISI©, EQ-5D and cost (reproduced from Figure 14 of CS)



Blue solid lines show main analysis (reference case). Gold solid lines show typical reference case pathway excluded from this evaluation due to lack of any serious safety concerns apparent in the data. Blue dotted lines show two potential routes to estimating productivity (non-reference case) either directly from the trial (SDS©) or mapped from ISI© (NHWS).
Tx=treatment; ISI=Insomnia Severity Index; HCRU=healthcare resource utilisation; QALY=quality-adjusted life year

We might also describe the model as a 'mediated' analysis. Arrows in the figure represent directly observed impacts. Treatment effect on ISI© was measured in the clinical trial programme (studies 301 and 303). ISI© impact on EQ-5D and HCRU was measured in an observational dataset (NHWS) and were estimated via 'mapping functions' and so the treatment effect on EQ-5D and HCRU was 'mediated' through ISI©. This type of model, although not common, does have precedent in the literature. Kuntz et al (37) describe this as a 'novel approach' in their 2002 paper looking at the relationship between FEV1 (the mediator) and symptom status in asthma trials. Briggs and colleagues (38) use a series of equations to estimate cost-effectiveness from the asthma 'GOAL' study which did not collect EQ-5D data. Another example from Briggs and colleagues (39) further developed the approach in a cost-effectiveness disease model for chronic obstructive pulmonary disease (COPD) (the 'GALAXY' model) where they described the modelling as a 'linked-equations cohort model'. We are not aware that any formal terminology has entered the lexicon, which was why we did not state the model form. However, the features of the model are described in full (it is a very simple model after all) and we are happy to elaborate below.

b) **Response:** We elaborate on the each of the ten-points raised below.

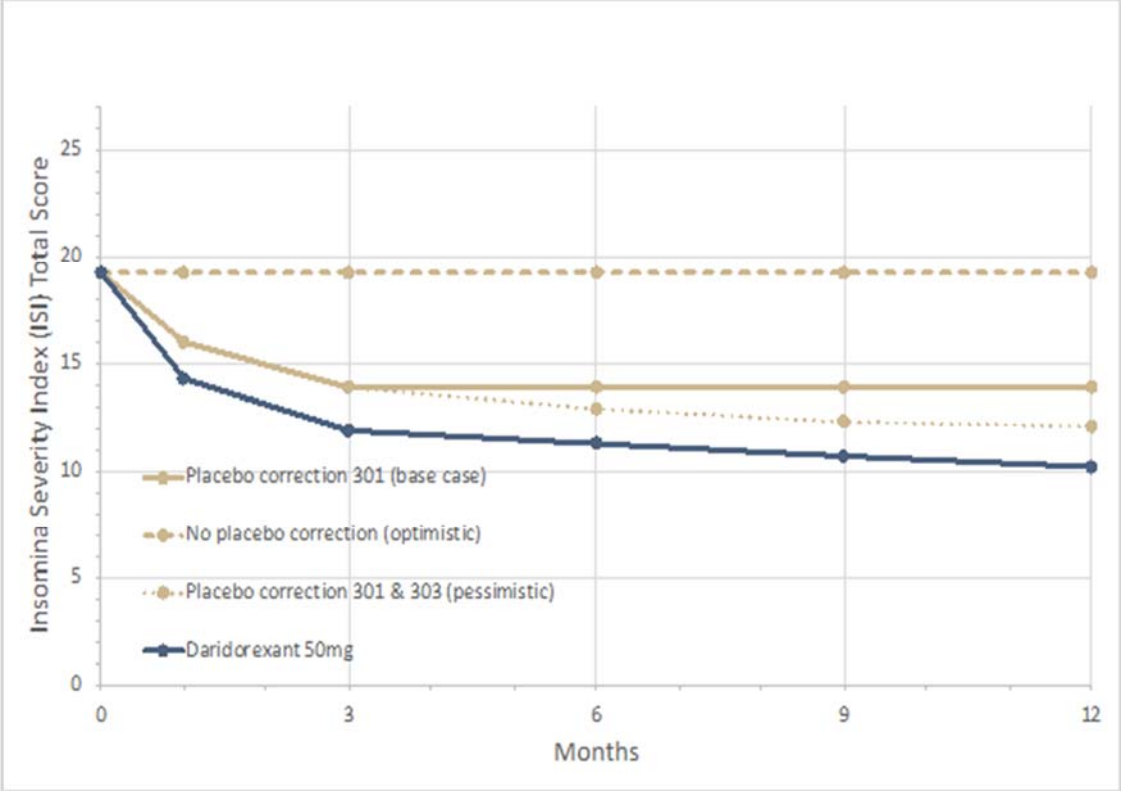
1. Time points relate to the following

Table 8: Timepoints of the study

Time-point in model (months)	Description
0	Baseline of study 301
1	Week 4 of study 301
3	Week 12 (end of follow-up) study 301
6	Week 14 (first post-baseline timepoint) study 303
9	Week 27 study 303
12	Week 40 study 303

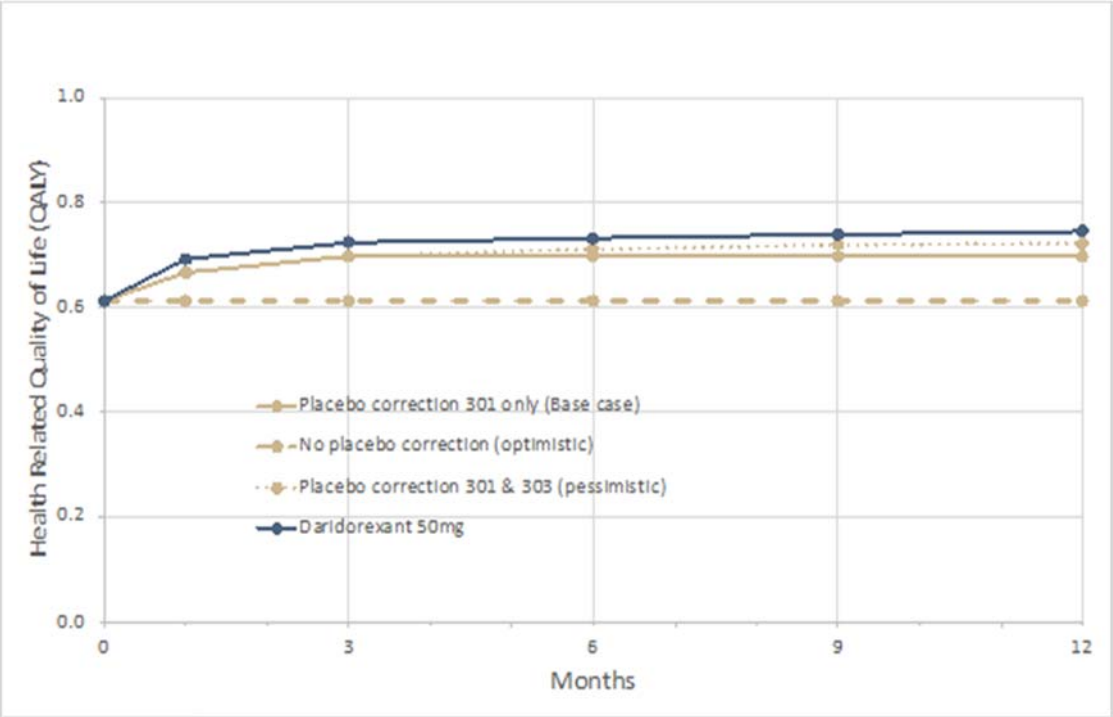
2. Since ISI© scores reflect a single observed value at the timepoints specific (coming from study 301 and 303 extension study) the mnISI calculation gives the average of the point at the beginning of the time window and the end of the time window as the average for that time period. If you consider the calculation of the area of a trapezoid then it should be clear that this is equivalent to an assumption of linear interpolation between the observed ISI© points. The observed ISI© at the given timepoints is presented in Figure 2 of the submission document (reproduced for your convenience below) as the dots in the figure with linear interpolation shown as the solid lines. Note the observed placebo group ISI© in the 303 study is the pessimistic scenario in the figure – the base case assumes no placebo correction as described in the CS and as elaborated in the response to the B9 priority question below.

Figure 2: Modelled trajectory of ISI© from phase III study 301 and 303 extension study showing base case, optimistic and pessimistic scenarios regarding placebo adjustment (reproduced from Figure 15 of CS)



3. For EQ-5D, the ISI© is first converted to an EQ-5D score using the mapping function to give EQ-5D at the given time point. Then mnEQ-5D is calculated as the average EQ-5D between time points – as before this is equivalent to an assumption of linear interpolation between timepoints. The conversion of the ISI© trajectories that comprise the model in Figure 3 of the submission were converted to EQ-5D scores and presented as a QALY profile in Figure 3 of the submission (reproduced for your convenience below).

Figure 3: Health Related Quality of Life utility profile of EQ-5D mapped from ISI© (reproduced from Figure 18 of CS)



EQ-5D=EuroQol-5D

4. GP visits, ER visits and IP stays were available from the same NHWS cross-sectional dataset that provided the mapping function between ISI© and EQ-5D. Three separate mapping functions were developed to describe the relationship between each of these available resource categories and ISI© score. Further details on these mapping functions were provided in the submission and are elaborated on in the responses to non-priority questions B15 and B16 below.

5. Number of administrations of 50mg of daridorexant per time period is assumed to be equal to all those remaining on treatment at the end of the previous time period. That is, we assume that the full cost of daridorexant is incurred for the full period, whereas the benefit in terms of improved health outcomes and reduced health care and productivity costs is dependent on those actively taking the medication. An implicit assumption is that if patients stop taking the drug, they do not return in the next period to fill a prescription.

6. The level of persistence (dropout) is given in row 71 of the <*Model | 1 Year*> worksheet. These persistence levels are equal to one minus the dropout rate from the beginning of the relevant time period. That is, these are 'further persistence/dropout rates' conditional on having persisted to the end of the previous period. These values are taken from the phase III 301 study and 303 extension study. These persistence/dropout values are implemented for active treatment only who are assumed to have values for the no treatment arm once they have dropped out from active treatment. The model uses the placebo arms as a proxy for no treatment – so persistence with 'no-treatment' is assumed to be 100% in the model. In terms of implementation: for health outcome, health care resource use and productivity losses, the model assumes a linear interpolation of drop outs between the rates observed at each time point. As described above, treatment cost is assumed to be 100% for the whole period regardless of dropout in order to mimic the potential waste that occurs when patients fulfil their prescription but do not complete the course of treatment.

7. The justification for adjusting for persistence/dropout is to make the model more 'realistic' recognising that not all patients will continue treatment. Assuming the full cost of treatment for patients starting the period but adjusting health outcomes and cost savings for persistence means that the net impact of including adjustment for persistence/dropout is to increase the ICER. So for example, the ICER unadjusted for persistence dropout is [redacted] per QALY. After the adjustment for persistence/dropout is made the ICER becomes [redacted] (the presented base case). Non-priority question B10 (d) makes the very reasonable observation that we argued that dropout in 303 is likely more representative of real-world persistence than the experimental phase III 301 study and requests a scenario where modelled dropout rates are based on 303. Dropout in 303 is [redacted] at 14 weeks, with [redacted] further dropout by week 27, and a further [redacted] by week 40. Using these figures to assume a dropout of [redacted] in month 1 (one third of the [redacted] rate), a further [redacted] over the next two months (so that dropout at month 3 is [redacted]), and then assuming the subsequent dropout rates at 6, 9 and 12 months are

[redacted] (the average of weeks 27 and 40 in study 303) yields an estimated ICER of [redacted] demonstrating that the model is relatively insensitive to small changes in persistence/dropout. Note that the ICER for those continuing on treatment at the end of 12 months remains at [redacted] whatever the underlying rate of dropout in the first year.

8. The cells in D48-I59 on the <*Model* | *1 Year*> worksheet refer to the incremental results assuming that everyone stays on treatment for the full time period. However, we wanted to adjust for persistence in the model which is what is done in D62 to I71 of the model. This is easily confirmed by changing the persistence figures in row 71 of the model to all read 100% so that both sections match exactly, and the cost-effectiveness improves to [redacted] for the base case. However, due to the selective attrition argument (see further elaboration in non-priority question B11 below) – we don't think this is a valid result which is why we did not use D48-I59.

9. Total costs are calculated as the sum of treatment costs, GP costs, ER costs, IP costs and (for non-reference case analysis only) productivity costs. For no treatment cost arm of the model, there are no treatment costs, but all the other cost categories are included. Since GP, ER, IP and productivity costs are a function of ISI© and ISI© scores are higher (worse) for the non-treated arm, the model does estimate cost savings attributable to daridorexant treatment. For health care resource costs (GP, ER and IP costs) these cost savings are estimated to be modest at just £20 per year (persistence adjusted) compared to the yearly cost of daridorexant of [redacted] (also persistence adjusted). However, the productivity cost savings are estimated to be substantial due to the substantial impact of insomnia on daytime functioning. When productivity costs are estimated from the SDS© collected in the trial, the cost savings are estimated to be [redacted] per year (persistence adjusted) which almost completely offsets the treatment cost. When the mapping function from the NHWS dataset is used based on the WPAI measure of productivity mapped to ISI, the persistence adjusted productivity cost

savings are estimated as [redacted] per year – more than offsetting the acquisition cost of treatment.

10. The model takes the opportunity to present many ICER values and we welcome the opportunity to explain these calculations in more detail. For each of three scenarios, labelled 'Evolving Cost-Effectiveness', 'First year cost-effectiveness' and 'Subsequent years cost-effectiveness', there are three types of ICER presented. Within each category, the estimated QALY is the same, only the costs differ. The first is an ICER based only on treatment cost. The second shows the ICER when health care resource use cost savings are netted off the treatment cost. The third ICER is when the total costs are presented net of health care resource use savings and gains in productivity. This last category of ICER is not a reference case analysis but is presented due to the important impacts of insomnia on daytime function which includes productivity.

    The evolving cost category (rows 75-77) presents the ICERs as they are calculated over time in the first year of the model – at each time point (see response to b(1) above). At the end of the year the outcomes are aggregated to give the average cost and average QALY over the first year of the model (persistence adjusted) and these are the figures presented in cells E80:82 (with E81 representing the base case ICER). Note how these values are the same as the yearly totals in J75:77. For those remaining on treatment at the end of the first year we can calculate an ICER in cells E85:88 based on the QALY and cost achieved at the last time point of the 12 month model (cells I62:71). The assumption here is one of 12 months costs and 12 months QALYs – though it should be clear that this is simply an 'on treatment' ICER and the time period is simply a scaling factor. This is the cost-effectiveness for remaining on treatment – however long that duration is.

c) **Response:** This question asks about the specific model characteristics – and mentions cycle time and half-cycle corrections as examples. As detailed in our response to (a) above, this is not a state-transition model. Subjects in the model do not reside in particular health states and they do not transition between those states. So half-cycle corrections and cycle times are not relevant. The first 12 months of the model is split into five time periods, however, as is described in the response to (b)(1) above and which relate directly to the observation time points in the phase III 301 study and the 303 extension study. The corresponding time periods are also clear from the Table in response to (b)(1) above. The first is one month, the second is two months and the final three are all three-month time periods adding up to 12-months in total.

d) **Response:** This question asks for more detail on the 'lifetime' cost-effectiveness model including estimation of parameters presented in Table 9 (reproduced below for convenience).

As described in the submission – the first year of the model is the short-term 12-month model that is presented as base case (which gave an ICER of [redacted] /QALY).

In subsequent years, the lifetime model takes the incremental costs of treatment and health benefit that comes from the subsequent year ICERs but only for those remaining on treatment.

The model could be described as a simple lifetable model. Only the incremental costs and benefits of being on treatment are modelled. Once a subject ceases treatment, either because they stop taking treatment or if they die then they cease to incur the incremental benefits of treatment as well as the incremental costs. The starting point is the 12-month base case model and the subsequent years cost-effectiveness with life table death rates applied to mimic the population survival but with an adjustment for the possible relationship between insomnia and survival. We assume that the application of lifetables is non-controversial and standard so we focus the

explanation here on how we adjusted for the possible effect of insomnia on survival.

Table 9: List of additional parameters to extend the short term-model to a lifetime model with names, values, description and the distribution used for the probabilistic analysis (reproduced from Table 63 of CS)

Name	Value	Distribution	Description
doRate	5%	beta	Annual Rate of dropout
btw67	1.01	lognormal	Relative risk of mortality for those getting less than 6hrs sleep
blw6hrs	1.04	lognormal	Relative risk of mortality for those getting 6-7hrs sleep
NT7plus	20%	Dirichlet	Proportion with sleep time of 7hrs plus without treatment
NT67	35%	Dirichlet	Proportion with sleep time 6-7hrs without treatment
NTblw6	46%	Dirichlet	Proportion with sleep time below 6hrs without treatment
D7plus	32%	Dirichlet	Proportion with sleep time of 7hrs plus on treatment
D67	33%	Dirichlet	Proportion with sleep time 6-7hrs on treatment
Dblw6	35%	Dirichlet	Proportion with sleep time below 6hrs on treatment
mnAge	50	NA	Average age at model entry
cDR	3.50%	NA	Annual discount rate for costs
oDR	3.50%	NA	Annual discount rate for QALYs

hr=hour; QALY=quality-adjusted life year; NA=not applicable

As stated in the CS: the evidence of an association of long-term health outcome and sleep was reviewed in B.1.3.2 Epidemiology. One of the few epidemiological studies that estimated a relationship between duration of sleep and mortality risk: Yin and colleagues conducted a meta-analysis of studies and reported the relative risk of low sleep duration (<6hrs per night) and 6-7hrs sleep duration on mortality risk as 1.04 and 1.01 respectively compared to a reference sleep duration of 7hrs or above (40). These are the relative risks that are given as parameters in Table 63 above. These can be combined with the estimated improvement in subjective total sleep (sTST) duration reported in study 301 (Table 14) to estimate the possible mortality benefits of daridorexant based on the average of 24 minutes increased sleep duration in study 301 (categorised to correspond to the definitions presented in Yin et al). The formal analysis showing that categorisation was not included in the CS but we have provided a Table

below (Table 10) to show how the data are categorised. Note that the parameters of the model are the percentages in each category but these are informed by the counts from study 301 which are also the parameters of the Dirichlet distribution used on the probabilistic sensitivity analysis.

We note that the mnAge parameter was listed as 50 in the CEM v2.23 with the initial submission but that the mean age at entry to the phase III 301 study was in fact 55. We have therefore updated the mean age parameter to be 55 in CEM v2.3.

*Table 10: Subject*ive total sleep time by category and by-arm for study 301 end of follow-up

	Numbers		**Percentage**	
Sleep time (hrs)	**PLA**	**D50mg**	**PLA**	**D50mg**
7+	61	99	20%	32%
6-7	107	103	35%	33%
<6	142	108	46%	35%
Total	310	310	100%	100%

Having clarified the general structure of the model and the calculation of the additional parameters, we now describe a step-by-step walk through of the calculations presented in each column of the lifetime model.

Column B is the year of the model and adding the starting age (now 55) gives the age in the model in column C.

Columns D and E show the incremental costs and incremental QALYs respectively drawn from the 12-month model as described above. The first year of the model (row 8 of the worksheet) refers exactly to the 12-month model. Therefore, the calculations described for each column below relate to year 2 onwards in the lifetime model (row 9 onwards in the worksheet).

Column G is the key calculation for the lifetime model as it provides an estimate of the additional QALYs that could occur through an improvement in sleep duration. We start by creating a weighted incremental relative risk of mortality between no treatment and treated cohorts based on the

estimated distributions across length of sleep categories reported in the table above.

For the Placebo arm this calculation is:

1 x 20% + 1.01 x 35% + 1.04 * 46% = 1.022

And for the daridorexant arm the calculation is:

1 x 20% + 1.01 x 35% + 1.04 * 46% = 1.017

These numbers represent the increased risk of mortality due to insomnia. Multiplying these numbers by the risk of death (which is looked up from standard ONS life tables given on the <*Ref* | *ONS Lifetable*> worksheet based on the current age in column C) and calculating the difference gives the incremental risk of death avoided from daridorexant treatment as a function of age.

To turn this into a life expectancy benefit, we mutiply this incremental risk of death by the life expectancy that would be lost if a death were to occur at that age which is agan looked up from the <*Ref* | *ONS Lifetable*> worksheet.

The final step is to quality adjust this life expectancy – we do this by applying the on-treatment estimate health related quality of life from the short-term model of 0.746 (cell I32).

As is apparent from column G, this possible additional mortality benefit of treatment is small relative to the estimated morbidity benefit estimated for the clinical studies. At age 55, the mortality benefit is less than 1% of the morbidity benefit rising to a possible 5% at older ages.

Column I estimate persistence with treatment over time. The starting point is the 55% persistence at the end of year one (though note this is not applied to the year one results as these are already persistence adjusted). Rather it is the starting point for persistence to which a further annual rate of persistence is applied (arbitrarily set to 5% per year in the model since we have no long-term data).

ERRATUM. In preparing the response to this question we spotted an error in our implementation of version 2.23 of the model which failed to adjust for death in the lifetime model as an added risk to discontinuation on top of the annual dropout rate. This risk of death is the same as is employed in column G and is looked up from the ONS Lifetable in the same way

Column J estimates the mean persistence in year 2 and onwards and could be considered to be similar to a half-cycle correction.

Columns M & N calculate the incremental costs and QALYs weighting by the persistence in column J and discounting based on the time in model (column B). Note that in column N the total QALY gain is estimated as the sum of the morbidity QALY from the short-term model (column E) and the mortality QALY (column G).

Columns O & P accumulate the persistence adjusted and discounted costs and QALYs from M & N respectively. The final column presents the evolution of the ICER from the base case estimate of [redacted] in the first year to the lifetime estimate of [redacted].

ERRATUM. The lifetime ICER figure quoted above is that produced by the corrected CEM v2.30. The original model and CS suggested a lifetime cost-effectiveness of [redacted]. So while we acknowledge there was an error in the submitted lifetime model, it turns out that it was not an important error for the interpretation of the results. The intuition for this is that it was only the death risk contribution to the persistence that was omitted – and this is a very small value in early years of the model. In later years when the death risk is much higher the model has already predicted that most subjects will have discontinued treatment.

B2. Priority question. Given the limited time horizon of the economic analyses, extrapolation of the trial data obtained from study 301 and 303 is not required.

a) Please justify the current model-based approach rather than a trial-based economic evaluation.

b) Please perform a trial-based economic evaluation (consistent with The Professional Society for Health Economics and Outcomes Research (ISPOR) Good Research Practices, see Ramsey et al. 2015 https://doi.org/10.1016/j.jval.2015.02.001) and provide both a detailed description of the methodology used (with justifications where appropriate) as well as the calculated results.

Response: We are pleased that the EAG recognises an extrapolation beyond 12-months of clinical data was not required (though we elaborate further in response to non-priority question B3 below). Given that the model is based on the three-month phase III 301 trial and the 40-week 303 extension study, then it may seem initially that a fully trial-based analysis should be possible.

a) **Response:** Neither the 301 study nor the 303 study collected health economic endpoints. Therefore, we do not believe a trial-based analysis is possible. We do, however, recognise that the model is something of a hybrid – being driven by the trial results through the mediating trial outcome ISI©.

b) **Response:** Given the lack of health economic outcomes, we contend the referenced ISPOR Task Force paper does not apply to the case presented here. It is clear from the abstract that the Task Force paper relates to studies that collect " ...clinical outcome measures, ... health resource use and health state utilities directly from study subjects." Later in the abstract the Task Force notes that "collection of economic data

should be fully integrated into the study." This was not the case for the 301 (and 303) studies which only collected clinical information. This is the reason we contend that this is a pathway model (or perhaps a hybrid trial-model). Nevertheless, we do a fair amount of trial-based analysis, particularly the seemingly unrelated regression of 301 which is covered in detail in the report and further elaborated on in response to priority question B8 below; as well as the Sheehan Disability Scale (SDS), a 301 and 303 trial questionnaire measuring directly the productivity losses, explained in detail in answer to question B17 (a).1.

B3. The CS base case analyses adopted a 12-month time horizon. The company stated that daridorexant has a quick onset and a short half-life. Therefore, the treatment benefit occurs when taking the drug and that treatment effect stops when treatment stops. The company expect this 12-month time horizon to capture relevant patient outcomes. Moreover, it is stated by the company that daridorexant has demonstrated safety for up to one year. Therefore, extrapolating beyond 12 months generates substantial uncertainty in the model.

a) Please clarify what proportion of patients is anticipated to be on daridorexant treatment 12 months after starting daridorexant treatment.

b) Please clarify what the cost implications are of daridorexant treatment after 12 months.

c) Please provide a detailed description of the inputs and assumptions used for the scenario analysis described in the CS, assuming that patients remain on treatment beyond 12 months (estimated ICER of [redacted] per QALY gained).

d) Please provide a scenario analysis, considering the population, input parameters and assumptions consistent with the CS base case, but only adding the estimated costs of daridorexant treatment after 12 months.

e) Please, given the above, provide further justifications for the 12-month time horizon, the implications and the uncertainty related to this assumption.

Response: We are happy to elaborate on these details.

a) **Response:** Row 62 of the <*Model* | *1 Year*> worksheet shows the cumulative persistance with treatment over the timepoints. By the end of the year, cell I63 shows that [redacted] of patients remain on treatment.

b) **Response:** The incremental cost implications of treatment (accounting for persistence) are shown in column J rows 64 to 71 of the <*Model* | *1 Year*> worksheet. Over the 12 months the incremental cost of daridorexant is estimated to be [redacted] (cell J64). There are estimated cost savings of £20 (cell J68) in health care resource costs and [redacted] (cell J71) of productivity costs. This leaves a net cost of treatment over 12 months (not shown in the model) of just [redacted] - £20 - [redacted] = [redacted].

c) **Response:** At the end of 12 months, cells I49 and I52 show the incremental utility as 0.048 and 3-month cost of daridorexant as [redacted] respectively. Cell I68 shows the estimated direct health care cost savings over the same 3-month period of £10. The ICER for those remaining on treatment at the end of the 12-months is therefore estimated as: 0.048 / ([redacted] -£10)*4 = [redacted]. This calculation is performed in cell E86.

d) **Response:** The requested scenario is already calculated in cell E85 of the model. The cost of daridorexant over 12 months is: [redacted] x 365 = [redacted] and the estimated QALY gain over 12 months for those remaining on treatment is the same 0.048 as given in (c) above. The ICER is therefore [redacted] as presented in cell E85.

e) **Response:** The rapid onset and short half-life of daridorexant means that the fundamental assumption of the model is that treatment benefit is obtained only for the days that subjects take the medication. For this reason we believe that the short-term model presented here, which covers the 12-month period of the clinical trial data, is sufficient to *"...reflect all important differences in costs or outcomes between the technologies being compared."* (NICE MPG36, 2022, paragraph 4.2.22). Further paragraph 4.2.25 of the same guidance states *"A time horizon shorter than a patient's lifetime could be justified if there is no differential mortality effect between technologies and the differences in costs and clinical outcomes relate to a relatively short period."* Since we are not modelling a mortality difference in the base case, and because of the quick onset and short half-life, we believe that focusing on the 12-months of the observed data for the model overs the most precise estimates of cost-effectiveness, since extrapolating beyond the observed data necessarily increases uncertainty.

Population

B4. Priority question. Currently, the CS is missing a description of the population that is included in the cost-effectiveness model.

a) Please provide a detailed description of the modelled population.

b) Please elaborate on how the modelled population matches with the NICE final scope, clinical practice in England and Wales as well as the trials used to inform the model.

c) The trials excluded *"subjects with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments"*.

Depending on the interpretation of these exclusion criteria, the trial may have excluded patients with a wide array of mental health problems. Considering that insomnia is frequently comorbid with other mental disorders, the exclusion criteria may have therefore excluded a significant percentage of patients that could usually be found in UK clinical practice for insomnia. Elaborate on the implications of the exclusion of those patients on the cost-effectiveness model and its generalizability to clinical practice.

a) **Response:** The modelled population is the same as the population enrolled into studies 301 (i.e., adults with insomnia disorder as per the DSM-5® criteria and with ISI© score ≥15). This is consistent with the population in the decision problem.

b) **Response:** As mentioned in the response to question A6 (c), the modelled population of adults with insomnia disorder as per the DSM-5® criteria is aligned with the NICE final scope. While the modelled population has a baseline ISI© score ≥15 (an inclusion criteria of study 301), this is not expected to result in a narrower population than that of the NICE final scope since ISI© <15 represents subthreshold insomnia (13).

c) **Response:** The company does not expect the exclusion of subjects *"with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments"* to impact the cost-effectiveness model and its generalisability to clinical practice. As mentioned in the response to question A11, the company acknowledges that a significant proportion of patients in clinical practice are likely to have comorbidities, including neuropsychiatric disorders. However, the underlying mechanisms of insomnia are thought to be the same in subjects with and without neuropsychiatric disorders, including depression.

Intervention and comparator

B5. Priority question. Study 301 includes dosages of 25 mg and 50 mg, Study 303 adds a dosage of 10 mg. According to CS section B.3.2.2 the economic model only includes the 50 mg dosage.

a) Please clarify factors that could determine the need for different dosages.

b) If authorization is only sought for the 50 mg dosage, see question A18, please reflect on the plausibility that all patients would receive the 50 mg dose in clinical practice in England and Wales.

c) Please provide an updated model and scenario analyses for patients receiving the 10 mg dosage and the 25 mg adjusting for both cost and effects of the different dosages.

a) **Response:** A lower dosage of daridorexant (25 mg once daily) is needed for patients with moderate hepatic impairment or where there is co-administration of moderate CYP3A4 inhibitors (see response to question A7(c)).

b) **Response:** MHRA authorisation will be sought for both 25 mg and 50 mg dosages of daridorexant. This is consistent with the approved marketing authorisation by EMA (32). It is anticipated that most patients will receive the 50 mg dose in clinical practice in England and Wales, except for a those mentioned in the response to question B5 (a).

c) **Response:** The 10 mg dosage of daridorexant will not be approved for use in England and Wales, assuming MHRA's label indication is consistent with that of EMA. As the 25 mg dosage is used only by patients with moderate hepatic impairment and those on concurrent moderate CYP3A4 inhibitors,

the costs and effects are anticipated to be the same as the population that would receive the 50 mg dosage, as both groups of patients are expected to achieve similar serum concentrations of daridorexant based on its pharmacokinetic profile. Therefore, it is not necessary to include scenario analyses with 10 mg and 25 mg dosages of daridorexant.

B6. Priority question. The NICE final scope specifies that established clinical management should be the comparator. The NICE CKS for insomnia mentions sleep hygiene, CBT-I, non-benzodiazepine hypnotic medication, zolpidem, zopiclone, benzodiazepines and melatonin. Figure 6 of the CS also highlights some of the comparators, including CBT-I, hypnotic drugs, Z drugs, and melatonin.

CS section B.3.2.2 argues that daridorexant is the first insomnia treatment with longer term data and that therefore no treatment is the appropriate comparator. However, data availability should not be a driving criterion when deciding upon the relevant comparators. Further, throughout the CS (for example Figure 1 of document A), the company presumably assumes that daridorexant will be given instead of other treatments (not in combination with other treatments).

a) Please provide an updated model with scenario analyses with fully incremental analyses including all relevant comparators (including sleep hygiene advice, CBT-I, melatonin, Z drugs, hypnotic drugs, benzodiazepines) in the model, including treatment specific costs and effects and provide the updated model (allowing the Evidence Assessment Group (EAG) to reproduce these analyses).

b) Please elaborate on the implications of assuming that daridorexant will be given instead of other treatments (not in combination with other treatments) and provide an updated model with scenario analyses where applicable to assess the impact of this assumption.

a) **Response:** The pharmacological treatment options for insomnia mentioned in the NICE CKS are neither recommended nor approved for long-term use and are therefore not considered as comparators for daridorexant in the CS. In addition, existing guidelines do not make specific recommendations for longer term treatment, and this was highlighted as an unmet need in the CS (B.1.3.6). This was discussed in detail during scoping, with feedback from clinical experts and patient groups, resulting in the removal of all comparators from the final scope and reconfirmed in the Decision Problem Meeting. The comments from clinical expert during the scoping meeting are reproduced here for reference: *"…At the workshop, attendees also discussed pharmacological treatments for insomnia. They explained that none of the currently approved pharmacological treatments are recommended for long-term use. Daridorexant is expected to be used to treat insomnia disorder, where symptoms last for more than 3 months per clinical trials. Therefore, the attendees agreed that none of the comparators listed in the draft scope are relevant. The scope has been updated to remove the comparators (41).*" Therefore, the company did not pursue a model including all relevant comparators since this is not relevant to the NICE final scope.

b) **Response:** The company does not anticipate daridorexant to be given in combination with other pharmacological treatments for insomnia according to its proposed positioning presented in B.1.3.6. In addition, subjects enrolled in studies 301 and 303 were not permitted to receive other treatments for insomnia other than CBTi. Therefore, there is insufficient evidence to support its use in combination with other pharmacological treatments for insomnia.

Effectiveness

B7. Priority question. The company mapped ISI scores, an exploratory trial outcome, to EQ-5D utility values and ISI score is thus the main driver of effectiveness estimated in the economic model.

a) How do other clinical endpoints included in the trials (wake after sleep onset (WASO), latency to persistent sleep (LPS), sTST, IDSIQ scores) compare to the ISI (for daridorexant compared with the placebo arm)?

b) Please provide further justification, per outcome, why the other clinical endpoints are not considered for mapping (in scenario analyses).

a) **Response:** We have compared these endpoints via an NNT analysis and presented the results in the response to question A20 (Table 5). The NNT analysis demonstrates that other clinical endpoints included in the trials (WASO, LPS, sTST, IDSIQ scores) are equivalent compared to the ISI (for daridorexant compared with the placebo arm). The mean (95% CI) NNT for an ISI responder (defined as a change from baseline of ≤ -7 points) at month-3 in daridorexant 50 mg is [redacted], compared to [redacted] for a sTST responder (defined as a change from baseline of ≥ 55 min), [redacted] for a LPS responder (defined as a LPS < 20 min at month-3), [redacted] for a WASO responder (defined as a LPS < 20 min at month-3).

b) **Response:** To the best of our knowledge, there is no dataset to map one of the other clinical outcomes (LPS, WASO via polysomnography; sTST via sleep diary) to the EQ-5D, alone or in conjunction with, the healthcare resource consumption (i.e., GP visit, ER visit, hospitalizations) and the work/daily activities productivity. The NHWS is a unique dataset as it allows mapping of the ISI© to these medico-economic outcomes. A single mapping avoids uncertainty associated with the mapping of different cohorts (e.g.,

mapping of one clinical outcome to EQ-5D based on a cohort, mapping of this same outcome to HCRU based on another cohort).

There is precedence in the use of similar mapping functions in sleep disorder (obstructive sleep apnoea). In August 2021 (NICE guideline NG202, "Obstructive sleep apnoea/hypopnoea syndrome and obesity hypoventilation syndrome in over 16s." Methods (2021) (42)), NICE released an economic guideline comparing the different types of continuous positive airway pressure machines (CPAP). The treatment effect of CPAP was measured on the Epworth Sleepiness Scale and mapped to the EQ-5D (43).

B8. Priority question. CS section B.2.9 explains how the company used seemingly unrelated regression for the cost effectiveness analysis. Section B.2.9.1 only describes the methods used for the analysis of the study period for study 301. Section B.2.9.2 describes what happens when patients discontinue or re-initiate treatment and does not illustrate the methods used for the study period of study 303.

In addition, to ascertain the quality of the application of this regression model, answer & report all questions asked by Kearns et al. guidelines for the use of regression models in health economic models (https://pubmed.ncbi.nlm.nih.gov/23807751/).

a) Please detail the methods and the rationale behind the methods used to obtain the seemingly unrelated regression model for the whole model period of 12 months.

b) Please justify the methods used (i.e. the seemingly unrelated regression model described in the previous sub question).

a) **Response:** In responding to this question we restate the rationale for using seemingly unrelated regression as part of the presented model (answer to part (b) of the question first). We are unable to present a regression of the whole 12-month period (part A) of the question above because the modelling team had access to patient-level data for study 301 only and the two studies, while having overlap, are not performed in exactly the same subjects. The 303 study results are taken from the aggregate data reported in the company clinical study report for the base case analysis with additional stratification for the subgroup analysis reported. Following the answer to question (b) we then provide a response to the 27-item checklist developed by Kearns et al and cited above.

b) **Response:** Seemingly unrelated regression (SUR) describes a series of linear equations that appear to be unrelated but are in fact related through the error term. In STATA SUR is implemented using the `sureg` command. When the regressions are 'unbalanced' (meaning they do not have the same explanatory variables) then it can be shown that there are efficiencies in estimation compared to running unrelated regressions of the same form. When the equations are 'balanced' (so the explanatory variables are the same) then there is no efficiency gain, but SUR does provide cross-equation correlations through a single covariance matrix.

The rationale for employing SUR for the estimation of the observed data in study 301 is as follows:

- SUR allows both the time points (month 1 and month 3) in study 301 to be modelled together in such a way that the regression equation exactly reproduces the observed values while capturing the correlation (through the error terms) between the two equations.

- This allows the appropriate sub-group analysis to be performed at each time point by including an interaction term between the sub-group variable (ISI severity of moderate or severe insomnia at the screening visit) and the treatment arm.

- Because SUR provides a cross-equation covariance matrix then the appropriate propagation of uncertainty can be performed in the probabilistic sensitivity analysis using the standard technique of Cholesky decomposition.

The use of SUR for cost-effectiveness data has been described in detail by Willan & colleagues (44). It is useful to note before going on to apply the Kearns check list, that SUR provided a framework for representing the trial results in a regression framework in a form that directly reproduces the observed values while allowing appropriate sub-group analysis with the relevant correlations matrices to allow propagation of uncertainty as described above. This is all described in the Willan & colleagues article cited above. The SUR is not used to develop a parsimonious statistical model and therefore much of the Kearns checklist is not directly relevant.

Application of the Kearn's checklist to the reported SUR of the 301 study.

1. Have the objectives of the analysis been stated?

The objective was to analyse the 301 study with a framework that would reproduce the observed outcomes while providing a framework for appropriate subgroup analysis and propagation of uncertainty.

2. Has the need for a de novo regression analysis been justified?

The regression presented is consistent with a standard by-arm analysis of the clinical trial data. The explanatory variables are the trial arm and the baseline ISI such that the base model reproduces exactly the estimated treatment effect in the trial.

3. Has the source of the data used been stated? This would include synopses of key study features such as socio-demographic/clinical characteristics and the data collection method.

Yes. Study 301 characteristics are extensively reported in the submission.

4. Has the total sample size available been reported?

Yes. A total of 557 subjects (out of 620 in the full analysis set) had ISI data at baseline, 1 month and 3 months.

5. Are sufficient explanations of all variables used provided?

Yes. The explanatory variables are the trial arm and the baseline ISI such that the base model reproduces exactly the estimated treatment effect in the trial. An indicator for severe insomnia at screening is used for the subgroup analysis.

6. Are sufficient numerical and/or graphical summaries provided?

Yes. Reported in the submission AND in the Excel model.

7. Has the quality of data (missing values, outliers, possible bias, etc.) been described?

Yes. A total of 557 subjects (out of 620 in the full analysis set) had ISI data at baseline, 1 month and 3 months. 10% of subjects had missing data.

8. Has the type/method of regression model(s) considered been stated/justified?

Yes. Seemingly unrelated regression (SUR).

9. Have any modelling assumptions been stated?

Yes. There are very few modelling assumptions as this is a standard OLS regression using simple treatment arm comparisons which means the regression itself is equivalent to undertaking a simple comparison of means in a by-arm comparison (for which one would not usually state assumptions).

10. Is a convincing rationale given for the inclusion of explanatory variables?

Yes. The explanatory variables are the trial arm and the baseline ISI such that the base model provides the estimated treatment effect in the trial. An indicator for severe insomnia at screening is used for the subgroup analysis.

11. Are sufficient details about the computational methods used provided?

The SUR was estimated using the `sureg` procedure in the statistical package STATA.

12. If more than one model was considered, has justification been given for why the preferred model has been selected?

Model selection was not attempted. The modelling form and the explanatory variables were chosen to replicate the trial analysis and provide a framework for appropriate subgroup analysis and propagation of uncertainty.

13. Has the choice of covariates been justified?

Yes. The explanatory variables are the trial arm and the baseline ISI such that the base model reproduces exactly the estimated treatment effect in the trial. An indicator for severe insomnia at screening is used for the subgroup analysis.

14. Is the sample size reported for every model presented?

Yes.

15. Has the handling of missing values (if any) been described?

Yes. Complete case analysis was performed. No imputation was considered necessary as the missing values were a small proportion of the total sample size (10%). Table 47 summarizes the demographic characteristics of subjects included in the analysis. Despite the missing information the characteristics remain similar to the full analysis set (Table 11).

16. Are the coefficient estimates provided?

Yes. See Table 48.

17. Are appropriate measures of uncertainty and significance provided?

Yes. See Table 48.

18. Are summary measures of goodness of fit presented?

No. Goodness of fit is not an issue. This is not an exercise in finding a parsimonious predictive model that provides a good fit to the data, rather the aim was simply to replicate the reported trial results within a framework that allowed the appropriate analysis of subgroups and associate uncertainty.

19. Are details of the results of a residual analysis provided?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

20. Has the model been validated on external (or quasiexternal) data?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

21. Is the plausibility of the modelled predictions and/or coefficients discussed?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

22. Are the results compared to the literature and/or other data?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

23. Has the method for handling parameter uncertainty been reported?

Yes. Cross-equation covariance matrices are used to propagate uncertainty in the probabilistic analysis of the decision model.

24. Is sufficient detail given for how parameter uncertainty was handled (e.g. if a variance–covariance matrix is used, is this available in some form?)

Yes. The covariance matrix is included in the Excel model and is used to calculate the Cholesky decomposition matrix to propagate the uncertainty.

25. Is parameter uncertainty appropriately reflected in the DAM?

Yes, the regression model is used to get the appropriate statistical quantities to propagate parameter uncertainty.

26. Has any structural (model) uncertainty been explored (in the DAM)?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

27. Have the model's limitations been discussed (and explored if possible)?

No. Not necessary/appropriate as the regression model mimics the standard by-arm analysis of the trial data.

B9. Priority question. In the base-case the company applies a placebo adjustment to adjust for the modelled patients receiving no treatment.

This adjustment relies on the assumption that the improvement in ISI scores is attributable to a placebo effect. It is unclear to the EAG whether the improvement in ISI scores is attributable to a placebo effect or whether patients with insomnia may also experience an improvement of their symptoms without treatment and that some of the observed improvements may stem from this 'natural improvement' (or regression to the mean).

a) Please clarify whether an improvement of symptoms over time may have been observed because of a natural improvement and not because of the placebo effect.

b) Please explain in detail how the placebo adjustment was implemented.

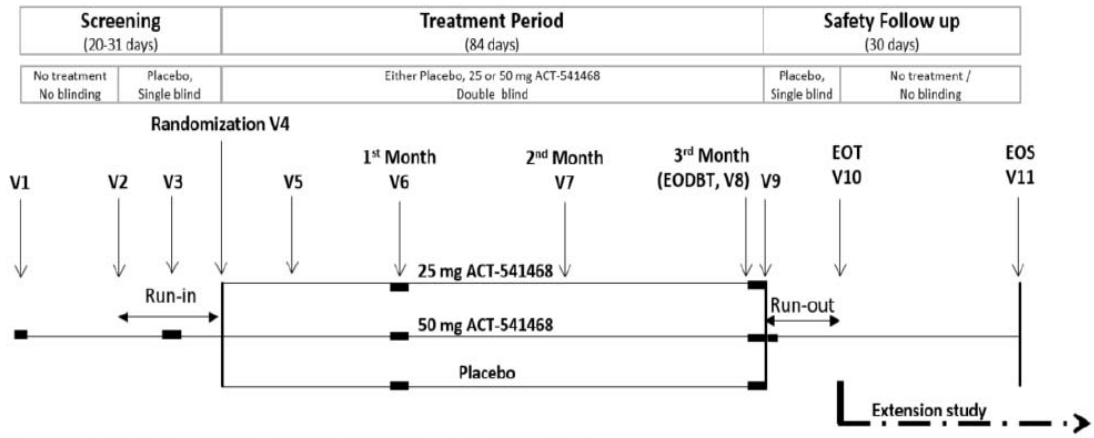
c) Please clarify per time-point what was done in the optimistic and pessimistic scenario analysis for the treatment and comparator arm.

d) Please conduct a scenario analysis without the placebo adjustment, i.e. without removing the placebo effect.

a) **Response:** It is reasonable to suppose that regression to the mean might be responsible for the observed placebo effect in the clinical studies reported. However, there are two main reasons why we consider that regression to the mean is not responsible for the observed placebo effect.

(i) The design of the study (reported in Figure 4 of the submission and reproduced here for convenience) shows that there was a screening phase. Visit 1 of the study recruited patients who scored ≥15 on the ISI© scale indicating they had at least moderate insomnia. However, randomisation occurred at Visit 4 some 20-31 days later. Regression to the mean does explain why the baseline ISI measure at randomisation is lower than ISI© at screening.

Figure 4: Design of study 301 (16) (Reproduced from Figure 7 of CS)

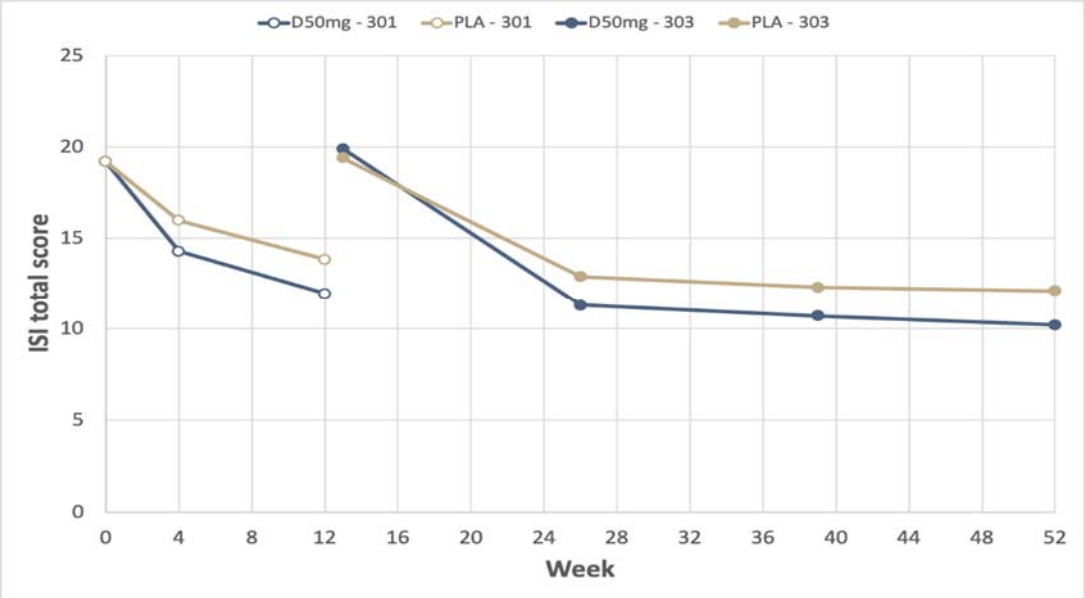


V5 and V11 were telephone calls; all other visits were at the site.
— = polysomnography nights; EODBT = end-of-double-blind-treatment; EOS = End-of-Study; EOT = End-of-Treatment; V = Visit.

(ii) Of course, regression to the mean could still explain the placebo effect post randomisation, however, the observed rebound effect between studies 301 and 303 suggests this is not the case. Although Figure 5 of the submission hints at this rebound effect, it is not quite apparent because the data are presented as change from baseline. An alternative presentation of these data is shown in

absolute terms in the figure below for ISI© scores of patients in 301 who continued on in the 303 study.

Figure 5: Change in ISI© scores from baseline to the end of extension study 303, all subjects included (reproduced from Figure 13 of CS)



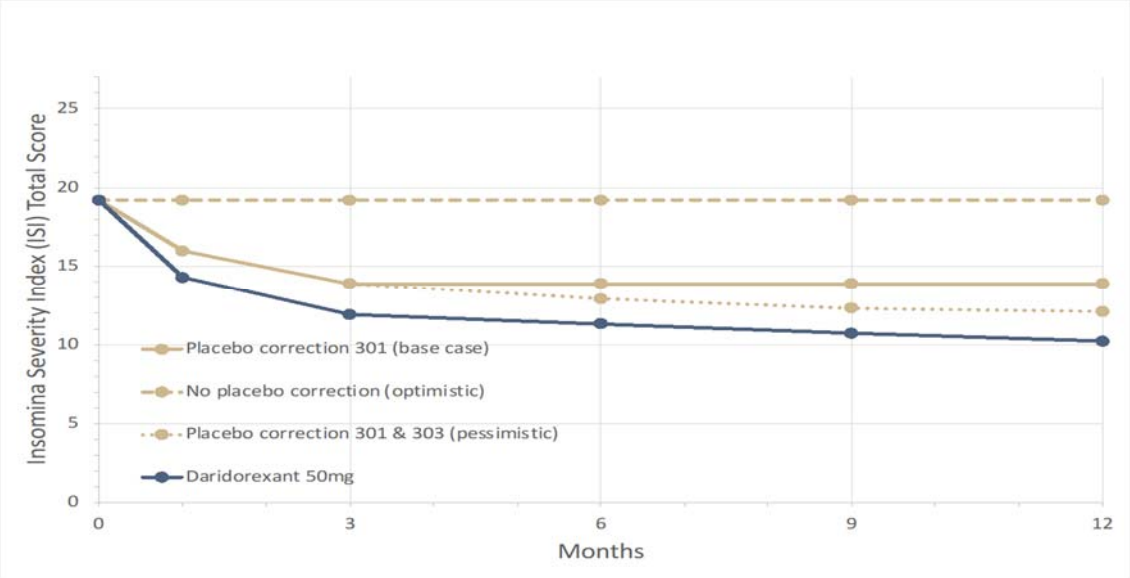
The figure clearly shows that following withdrawal of treatment during the 'run-out' week after the 301 study, those patients who continued into the 303 extension study had returned to the baseline values they were experiencing prior to the 301 study commencing. If regression to the mean were responsible for even some of the treatment effect, we would expect the rebound to return to something less than the original baseline.

b) **Response:** The placebo adjustment is simply the difference between the treated arm of the studies and the placebo arm of the study i.e., the placebo adjustment used in the model is the standard ITT estimate from the phase III 301 study.

c) **Response:** Figure 6, reproduced from the submission presented, for each time point the assumptions concerning the ISI© trajectories for the base case analysis and the optimistic and pessimistic (best/ worst case) scenario analyses. The figure is reproduced below for convenience. Recall that the first

3 months of the trajectory comes from the 301 data and the subsequent 9 months comes from the 303 study.

Figure 6: Modelled trajectory of ISI© from phase III study 301 and 303 extension study showing base case, optimistic and pessimistic scenarios regarding placebo adjustment (reproduced from Figure 15 of CS)



As Figure 6 shows, the base case assumes a placebo adjustment from the 301 study but caps the placebo adjustment at that 3-month point. By contrast the dashed line shows the best-case scenario where all the improvement from baseline is attributed to treatment (that is no placebo adjustment) - this generates an estimated ICER of [redacted] (see Table 60 of CS). The worst case scenario assumes that there is a continued placebo adjustment into the period of the 303 extension study ignoring the evidence that the continued improvement in ISI© is due to selective attrition (see response to non-priority question B11 below) - this generates an estimate ICER of [redacted] (see Table 59 of CS).

d) **Response:** We are confused by this part of the question and assume that there is a typo since the questions asks us to present a scenario without the placebo adjustment but clarifies that to be without removing the placebo effect. Nevertheless, we trust that our response to (c) above makes it clear that the

best and worst case scenarios already presented in the submission cover both no placebo adjustment and full placebo adjustment.

B10. The CS states that treatment discontinuation was modelled based on observed discontinuation rates with reference to CS Figure 16. It is unclear to the EAG how discontinuation rates were estimated and incorporated in the (input parameters of the) model.

a) Please detail when and how many patients were modelled to discontinue treatment.

b) Please detail how the discontinuation rate was estimated and incorporated in the (input parameters of the) model.

c) The CS states that study 303 likely reflects clinical practice more accurately. However, according to the EAG's understanding of Figure 16, discontinuation rates from study 301 and study 303 are applied in the model. Please justify applying discontinuation rates from study 301 when these discontinuation rates are less likely to reflect clinical practice.

d) Please conduct a scenario analysis applying discontinuation rates from study 303 throughout the whole model.

We are happy to clarify how the data in Figure 16 were calculated.

a) **Response:** In the Table 11 we show the numbers of patients continuing treatment for each time point along with a reminder of which study is being used.

b) **Response:** The table also shows the conditional probabilities of dropout which formed the basis of Figure 16 of the CS.

Table 11: Conditional dropout rates in study 301 and 303

		Numbers in study		Conditional dropout	
Study week	Model month	D50mg	PLA	D50mg	PLA
301 baseline	0	310	310	0%	0%
301 week 4	1	299	297	4%	4%
301 week 12	3	283	281	5%	5%
303 baseline	-	[redacted]	[redacted]	-	-
303 week 14	6	[redacted]	[redacted]	[redacted]	[redacted]
303 week 27	9	[redacted]	[redacted]	[redacted]	[redacted]
303 week 40	12	[redacted]	[redacted]	[redacted]	[redacted]

c) **Response:** Our justification for using study 301 discontinuation rates is simply that we chose to use observed discontinuation rates throughout. However, we agree that 303 discontinuation is likely to represent real life since 303 was an open label extension study not a phase III explanatory trial.

d) **Response:** We have conducted the requested scenario analysis using drop out rates based on 303 (note this scenario is also described in the response to B1 (7) above). Drop out in 303 is [redacted] at 14 weeks, with [redacted] further dropout by week 27, and a further [redacted] by week 40. Using these figures to assume a dropout of [redacted] in month 1 (one third of the [redacted] rate), a further [redacted] over the next two months (so that dropout at month 3 is [redacted]), and then assuming the subsequent dropout rates at 6, 9 and 12 months are [redacted] (the average of weeks 27 and 40 in study 303) yields an estimated ICER of [redacted] demonstrating that the model is relatively insensitive to small changes in persistence/dropout. Note that the ICER for those continuing on treatment at the end of 12 months remains at [redacted] whatever the underlying rate of dropout in the first year.

B11. An assumption is made for the modelling of the period of study 303: The improvements in ISI were not due to an increasing effect of treatment but due to selective attrition. For the no treatment group, a placebo adjustment was modelled arguing that study 303 presented evidence of selective attrition. Given the nature of study 303 as a long-term extension study, selective attrition could

also happen in the treatment arm, where only patients who do not experience a treatment effect drop out. It is unclear to the EAG how Figure 13 in the CS shows that selective attrition happened in the placebo arm but not in the treatment arm.

Please explain why the effect of selective attrition was only assumed for the placebo arm but not for the treatment arm.

Response: We think there is some confusion here. We use Figure 13 to justify that there is selective attrition in both the placebo and treated groups of 303. This is apparent in the lower change in ISI© from baseline from those that drop out from either arm than for those that continue in either arm.

When we take the data into the model, we use the selective attrition argument for the treatment arm only (not as stated above for the placebo arm only). The reason is that the placebo arm is a proxy for no treatment, but it makes no logical sense within the model for patients to be dropping out from no treatment. We hope this clarifies the assumptions in the model.

Adverse events (AEs)

B12. According to the CS, during the double-blind study period in study 301 and study 303, 37.7% and 38% of the subjects reported at least one treatment-emergent AE in the daridorexant 50 mg arm respectively, while 1% and 5.1% of the participants experienced at least one treatment-emergent serious AE. However, the AEs reported in both studies were of mild or moderate intensity and expected to not have a significant impact on the HRQoL and patient costs, and therefore, it was not included in the model.

Please provide an updated cost-effectiveness model and scenario analyses incorporating all AEs from study 301 and study 303 as well as the impact on estimated costs and effects.

Response: We stand by our assertion that daridorexant has excellent safety data. Indeed, the label indicates that the regulator agrees that side effects are minor and not substantially different from placebo. While the EAG is correct that Table 18 shows that 37.7% of the treatment group reported at least one adverse event, it is also apparent that 34% of the placebo group also reported at least one adverse event. Although it could be argued that there is a numerical increase of 3.7% in the treatment group it should also be clear to the EAG that the listed adverse events are minor such that they are unlikely to have any HRQoL or resource use consequences. Regulatory studies require a cautious and broad reporting of possible adverse events.

We believe that it is much more common, within the context of a health economic model for NICE submission, that quantitative inclusion of adverse event data be limited to serious adverse events (SAEs) that are much more likely to have HRQoL and HCRU consequences. Table 19 of the CS shows potentially treatment related serious adverse events of any type are fewer numerically in the treatment arm (1%) of the 301 trial than in the placebo group (2.3%) and are of very low overall frequency (<2%). After adjudication for adverse events of special interest the overall frequency falls to below 0.5% as shown in Table 20 of the submission.

The inclusion of all AEs in the model will not have a substantial impact on the ICER. This is illustrated in the following scenario. The mild nature of the reported adverse events means that a utility decrement of 0.2 would be a large decrement, suppose that the duration of the event is 5 days (which is surely an overestimate) and that the AE is treated by a simple trip to the pharmacy for an over-the-counter remedy that costs £5. The QALY loss would be 0.2 x 5 / 365 = 0.0038 per event multiplied by the 3.7% additional events in the treated group giving a net QALY loss of 0.00014. Subtracting this from the base case results

and adding £5 x 3.7% = 18.5p to the incremental costs moves the reported base case ICER from [redacted] to [redacted].

Quality of life

B13. A generalised linear model was used to create a mapping function from the cross-sectional National Health and Wellness Survey (NHWS) survey to derive EQ-5D utilities from ISI© scores reported in studies 301 and 303. However, limited information is provided for the EAG to consider the (development of the) mapping function.

a) The population used for developing the mapping algorithm (from the NHWS survey) was broader than the trial population with regards to psychological problems. Please elaborate on the differences between the trial population and the population that was used for the development of the mapping algorithm and the potential implications these differences may have on the estimated utility values used in the economic model.

b) Please elaborate on the conceptual overlap between the ISI and EQ-5D instruments.

c) Please provide statistics regarding the correlation between the elements of the ISI and EQ-5D instruments.

d) Please justify, considering the responses to the preceding sub questions, that it is appropriate to map the EQ-5D utilities from ISI© scores.

e) Please provide further justification why other outcomes measured in the pivotal trials are not suitable for mapping to the EQ-5D.

f) Please consider the ISPOR Good Practices for mapping studies (Wailoo et al. 2017 https://doi.org/10.1016/j.jval.2016.11.006) and provide detailed responses to all aspects/considerations mentioned in Tables 1, 2 and 3 of this paper.

g) Please provide an updated economic model and scenario analyses, incorporating an updated mapping function considering the ISPOR Good Practices for mapping studies.

h) Please provide an updated economic model and scenario analyses, incorporating published mapping functions between ISI and the EQ-5D (including the mapping function reported by Gu et al. 2011 https://doi.org/10.1186/1477-7525-9-119).

i) Please provide an updated economic model and scenario analyses, incorporated a re-estimated mapping function, including relevant covariates (including at least age and gender), providing detailed responses to the ISPOR Good Practices for mapping studies, while using the following model types (mentioned in Gray et al. 2018 https://doi.org/10.1016/j.jval.2017.09.017):

    1) Adjusted Limited Dependent Variable Mixture Model (ALDVMM)

    2) Censored Least Absolute Deviations (CLAD)

j) Please report the root mean squared errors (RMSE) and the mean absolute errors (MAE) for all mapping functions.

We welcome the opportunity to elaborate on the mapping model.

The Table 12 below is taken from the submitted Excel model <*Ref* | *NHWS utility*> and shows the fitted model coefficients as well as their baseline values in the NHWS data set.

Table 12: Fitted model coefficients as well as their baseline values from the NHWS dataset

	Coefficients	Standard errors	NHWS means
const	-1.339	0.023	
female	-0.053	0.011	0.673
pain	0.250	0.011	0.417
degree_4years	-0.059	0.011	0.441
married	-0.028	0.010	0.552
employed	-0.117	0.013	0.590
retired	-0.044	0.017	0.172
current_smoker	0.118	0.013	0.242
former_smoker	0.049	0.012	0.265
low_moderate_drinker	-0.080	0.012	0.612
heavy_drinker	-0.042	0.018	0.126
Combined_DP_PTS_AX	0.289	0.011	0.409
ISI_Score	0.249	0.006	0
CCI	0.075	0.005	0
BMI_R	0.068	0.005	0
treated	0.037	0.016	1
treated & ISI_Score	-0.055	0.014	-0.055
Not_Country_UK	-0.061	0.016	0

Example of calculation, if the utility score to predict is for an ISI summary total score is equal to 11:

$$f(x) = 1 - exp\left(-0.05349 * \left(\frac{11992}{17955}\right) - 0.02758 * \left(\frac{9850}{17955}\right) - 0.05933 * \left(\frac{8095}{17955}\right) - 0.11709 * \left(\frac{10534}{17955}\right) - 0.04402 * \left(\frac{3142}{17955}\right) + 0.118266 * \left(\frac{4377}{17955}\right) + 0.049383 * \left(\frac{4782}{17955}\right) - 0.04209 * \left(\frac{2273}{17955}\right) - 0.08037 * \left(\frac{10989}{17955}\right) + 0.010669 * \left(\frac{26.67 - 26.67}{26.67}\right) + 0.102955 * \left(\frac{0.3 - 0.3}{0.3}\right) + 0.289408 * \left(\frac{7380}{17955}\right) + 0.249725 * \left(\frac{7506}{17955}\right) + 0.036965 * \left(\frac{2432}{17955}\right) + 0.046741 * (\mathbf{11} - 12.12) - 0.01028 * \left(\frac{2432}{17955} * (\mathbf{11} - 12.12)\right) - 0.06051 * \left(\frac{17955 - 2128}{17955}\right) - 1.33939\right)$$

$$= 1 - exp\left(0.046741 * (\mathbf{11} - 12.12) - 0.01028 * \left(\frac{2432}{17955} * (\mathbf{11} - 12.12)\right) - 1.3309\right) = 0.7488$$

a) **Response:** There are not too many overlaps in terms of the characteristics of patients from the trial and the subjects in the NHWS dataset. Age was not

included in the model because it was insignificant, though the average age of the NHWS sample was younger at 46 years compared to 55 years among those recruited to the 301 study. The proportion of women was similar at 67% in NHWS compared to 65% in the 301 study as was BMI at 26.7 in the NHWS study compared to 26.3 in the 301 study. Probably the most important difference was the ISI© score which had a mean of 12.6 (subclinical) in NHWS compared to 19.2 (moderate) in the 301 study. We do not think that there is a conceptual problem with developing the mapping function on a broader range of severity than the clinical trial. Indeed, it could be argued that this is a positive attribute since a broader range of ISI© and EQ5D values should result in a more robust mapping algorithm.

b) **Response:** We did comment on the conceptual overlap between ISI© and EQ-5D in Section B.3.13 of the submission. The mapping algorithm above has demonstrated that insomnia disorder as measured by ISI© does correlate with EQ-5D and was suitable to estimate the QALYs presented in this submission. Nevertheless, it is also very plausible that EQ-5D does not fully capture the impact of insomnia disorder on HRQoL. It has long been understood that EQ-5D may miss important dimension of HRQoL for some conditions – past research has explored the potential use of 'bolt-on' dimensions to capture missing dimensions. One of the most popular candidates for a bolt-on is fatigue, as fatigue is a feature of many health conditions including insomnia disorder (45). Perneger and Courvoisier examined possible missing dimensions from EQ-5D and identified separately fatigue/ energy and sleep as two dimensions that are poorly represented (46). Therefore, we believe it is reasonable to consider that the QALY estimates presented in this submission are an underestimate of the benefits of daridorexant on HRQoL.

c) **Response:** The model presented is a GLM and so strictly speaking an R2 statistic (where R2 is the square of the correlation) is not estimated. However, in an OLS model the R2 statistic was 35% suggesting the correlation between ISI© and EQ-5D after adjustment for covariates in the model is approximately sqrt(0.35)=60%. We consider this a good correlation.

Below, the non-adjusted GLM models summary statistics show that the ISI© 28-item (7 questions with 4 levels, 0 being the reference level) model is comparable to the ISI© total score (R2: 0.194 vs. 0.179, BIC: -11,567.3 vs. -11,652.0, respectively; The BIC penalising more heavily for complex models). As per their coefficients, levels of questions related to daytime functioning symptoms (Q5 to Q7: interference with daily functioning, worry about sleep, and sleep dissatisfaction) correlated better than night-time symptoms questions (Q1 to Q4: sleep onset and sleep maintenance difficulties (both nocturnal and early morning awakenings), satisfaction with current sleep pattern).

ISI© Total Score versus ISI© Domains

Table 13: Numerical results for the GLM models based on the individual questions and the ISI total score.

	MSE	R2	Log-likelihood	AIC	BIC
ISI© questions	0.040865	0.194164	5925.7	-11793.3	-11567.3
ISI© total score	0.041620	0.179260	5835.8	-11667.6	-11652.0

Table 14: Coefficients and p-values of the GLM model based on the individual ISI questions.

	const.	Q1_1	Q1_2	Q1_3	Q1_4	Q2_1	Q2_2	Q2_3	Q2_4	Q3_1	Q3_2	Q3_3
coeff.	-1.8388	0.0084	0.0563	0.1309	0.1917	0.0671	0.1343	0.2055	0.2773	-0.0156	-0.0021	0.0165
p-value	0.000	0.644	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.320	0.895	0.390

	Q3_4	Q4_1	Q4_2	Q4_3	Q4_4	Q5_1	Q5_2	Q5_3	Q5_4	Q6_1	Q6_2	Q6_3
coeff.	-0.0057	-0.1036	-0.1093	-0.1156	-0.0659	0.2067	0.3123	0.4334	0.5467	0.1022	0.1596	0.2501
p-value	0.831	0.007	0.003	0.002	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000

	Q6_4	Q7_1	Q7_2	Q7_3	Q7_4
coeff.	0.2349	0.057	0.1117	0.1644	0.2153
p-value	0.000	0.001	0.000	0.000	0.000

Table 15: Coefficients and p-values of the GLM model based on the ISI total score.

	constant	ISI© score
coefficient	-1.2951	0.0579
p-value	0.000	0.000

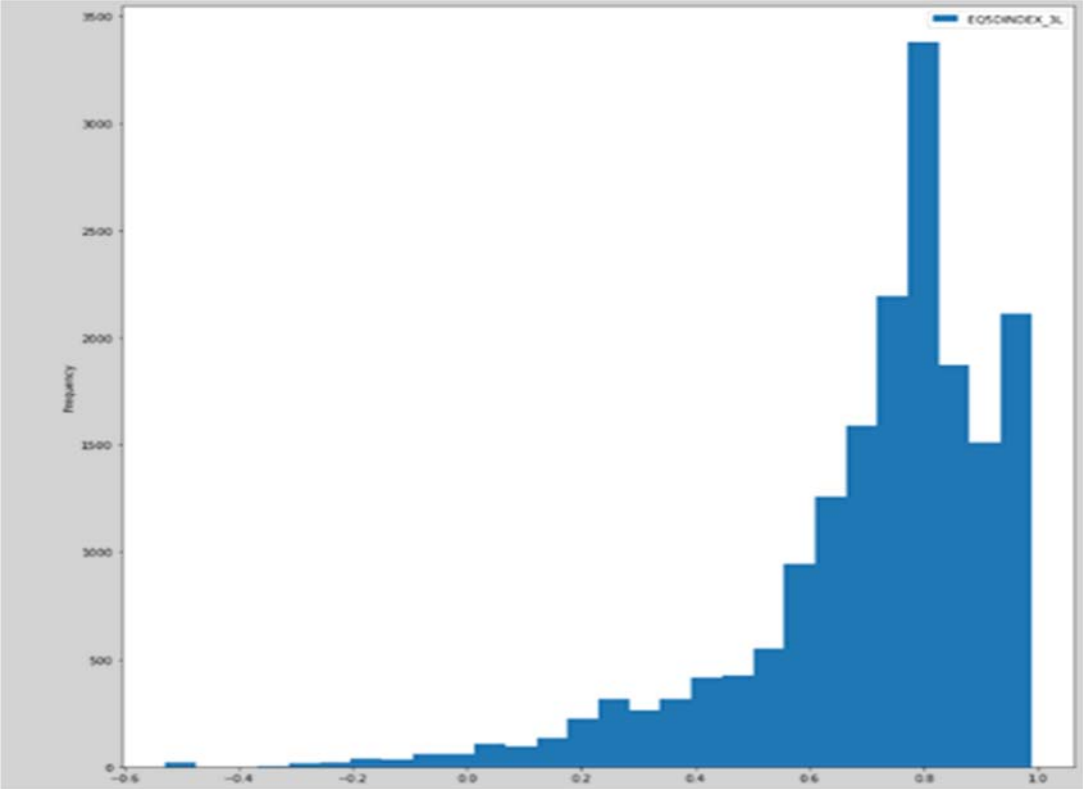
d) **Response:** We have conducted a mapping exercise, consistent with good practice guidance on mapping studies, to show that ISI© can be mapped to EQ5D. We consider the fit to be good, but nevertheless an underestimate since there is some evidence that EQ-5D does not cover all relevant domains for insomnia-related HRQoL due to known deficiencies of EQ-5D in capturing fatigue – we therefore suggest that the QALY estimation presented in the model is likely an underestimate of the true health benefits of treating insomnia.

e) **Response:** It is not that we consider other measures unsuitable for mapping to EQ-5D, simply that there are no available data sources to use to estimate a mapping.

f) **Response:** We have reviewed the ISPOR Good Practices Guide cited, with particular reference to the items presented in Table 1 (pre-modelling considerations), Table 2 (modelling and data analysis) and Table 3 (reporting). We believe that the development of the mapping function is in alignment with this Good Practices guide, although we acknowledge within the CS we did not fully report on a number of aspects that were recommended. In particular, although we included Figure 7 on the distribution of the EQ-5D (reproduced below for convenience), we did not link that to the reason for the model choice clearly enough. Also, it was recommended to include a figure on the fitted

versus observed estimates across a severity range. We now elaborate on these points below.

Figure 7: Distribution of the EQ-5D tariff scores in the NHWS dataset (reproduced from Figure 17 of CS)

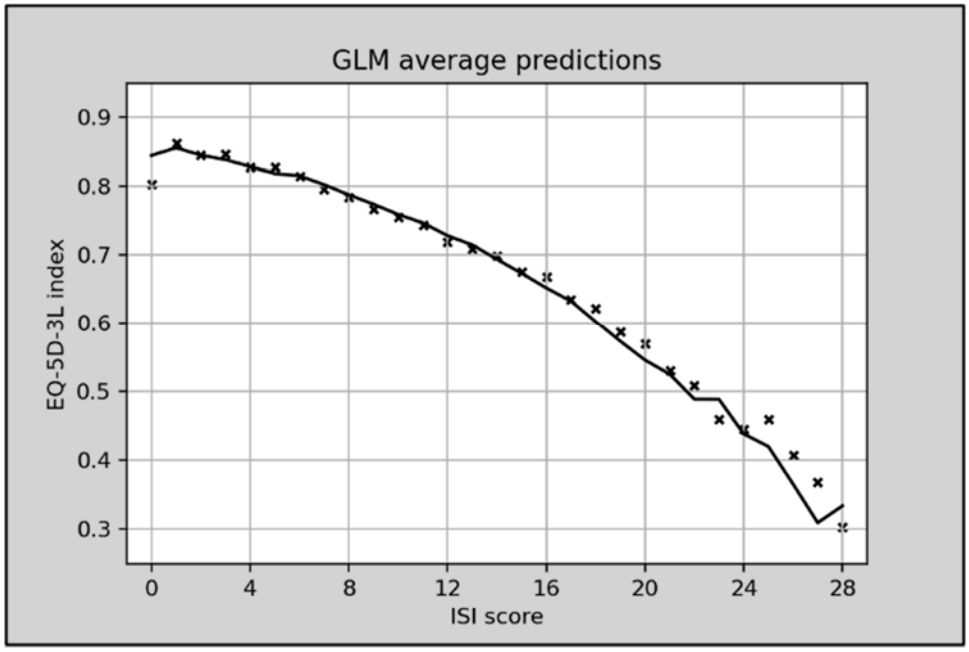


The distribution clearly shows that EQ5D values have a left skew distribution. The Task Force report recommends that model selection is based on "...the most straightforward statistical model type whose assumptions are compatible with the target utility instrument" and that analysts should use "...a plot of the distribution of the utility data to help inform that choice." Based on Figure 17 of the CS we chose to model disutility =1 - utility as described in the submission as it is a simple transformation that renders the data right skew instead of left skew with a Gamma-Log GLM to handle the skewness in the data. This is a simple model, consistent with the data distribution and consistent with previous mapping studies including the Gu et al paper referred to in part (h).

The Task Force Report also recommends to provide "...information on fit conditional on disease severity as measured by the clinical outcome measure(s). A plot of mean predicted versus mean observed utility conditional

on the clinical variable(s) should be included". We are pleased to provide this additional figure below showing an excellent fit of the model. On average the model does not show any bias in the extreme values of ISI© (as the Gray et al citation suggests is a common problem with 'standard' methods) – though of course the very extreme parts of the estimated figure are more noisy due to fewer observations.

Figure 8: The GLM model average predictions on the dataset (solid line) and true average values (crosses).



g) **Response:** Since we consider that the mapping exercise conforms to the ISPOR best practice we have not updated our preferred model and so the base case mapping function remains.

h) **Response:** Since the Gu et al mapping function is the only other mapping function between ISI© and EQ-5D that we are aware of, this is the only change we have made to the model (47). We do not consider Gu et al to represent an appropriate base case because their mapping function was conducted in a US sample of subjects using an out-of-date and US-specific tariff. Nevertheless, we have added the functionality to v2.3 to switch to the Gu et al algorithm as an alternative to our NHWS algorithm. Using the Gu et al algorithm results in an ICER of [redacted] compared to our preferred base case of [redacted],

nevertheless the results are relatively insensitive to the choice of mapping algorithm.

i) **Response:** Unfortunately, as the data set on which the mapping was conducted is held by a third part vendor, there simply is not time, within the two-weeks of responding to these queries, to re-estimate the mapping function. Furthermore, having reviewed the NICE guidance, the ISPOR Task Force Report, and the cited article, we respectfully suggest that request to re-estimate the mapping function is unnecessary. Nowhere does the Task Force report suggest that ALDVMM and CLAD models are gold standard methods that have to be considered. Indeed the Task Force Report specifically calls for simple methods to be employed where possible. The cited Gray et al paper is a cutting-edge methodological piece, and it is common that cutting-edge methodology is not the basis of guidance which tends to focus on more standard methods. But even as a cutting-edge piece they only compared the ALDVMM model to the standard linear model not to the GLM on disutility which given the typical shape of utility data is the preferred approach. The referenced article Kaambwa et al within the Gray et al paper suggests GLM outperforms standard OLS models. Neither does the Gray et al paper employ a CLAD model – the reference to CLAD models is through there reference to the Kaambwa et al paper. We have a theoretical objection to the use of CLAD estimators since they favour median estimates over the preference in health economic evaluations for expected (mean) values. These is clearly apparent in Table 3 of Gray et al (who reproduce Kaambwa et al's results) where the RMSE of the CLAD estimator is worse than either the OLS or GLM models but the MAE is better. Since RMSE is consistent with mean estimation, performance on RMSE is preferred to performance on MAE.

j) **Response:** Since we did not report any new mapping functions we did not report RMSE or MAE for the new models. Although we acknowledge that RMSE and MEA are often reported together (Gray et al and Kaambwa et al being prime examples) we question the need for MAE. RMSE is clearly preferred for two reasons: first it favours models that correctly predict expected values whereas MAE favours median estimators; second, (R)MSE combines both bias and

variance, so it is possible to compare biased and non-based estimators. There is also a practical challenge to presenting both together – since they are both reported on the same scale then the untrained reader might be tempted to compare the magnitude of RMSE and MAE – but they are not strictly comparable since by definition MAE will always have a smaller value than RMSE.

The RMSE for our preferred algorithm was 0.188 which is marginally higher than a model that included age and whether subjects had self-reported that they had received an insomnia diagnosis from a clinician. However, these two covariates were insignificant, and we chose our favoured model based on the AIC which was lower for the more parsimonious model. This is another factor to keep in mind when comparing RMSE – that there is no adjustment for the number of parameters included in the model.

B14. In the CS, the company mentioned previous evaluations by NICE, including technology appraisal (TA) 77 and MTG70, that focussed on technologies indicated for short-term management of insomnia symptoms.

Please provide an updated model and scenario analyses using utilities derived from other relevant TAs and provide a justification of how these compare to the utility values currently used.

Response: TA77 summarised guidance on short-term use of so-called 'Z-drugs' to treat insomnia (48). However, the report noted that "*No comparative data on the health related quality of life associated with Z-drugs and benzodiazepines using generic health status measures were identified, and there was no evidence to link the clinical endpoints from the trials with quality of life.*"

Similarly in MTG70 (21), which looked at the Sleepio® app as a form of app-based CBTi the company only submitted a cost-model claiming equivalent outcomes to face-to-face CBTi. No utility data was used to inform the guidance.

Costs and resource use

B15. Table 53 of the CS provides the treatment cost for daridorexant (50 mg) and a discontinuation adjustment for the annual treatment cost. No costs are reported for the no treatment comparator.

a) Please provide a step-by-step explanation as to how the overall treatment costs were calculated. Also include this level of detail for all comparators requested in question B6.

b) Please provide details regarding the discontinuation-adjustment calculation for daridorexant costs and include discontinuation-adjustment for all comparators requested in question B6.

c) The model includes costs for the no treatment group. Please justify why no costs are assumed for the no treatment group in Table 53 of the CS and provide details regarding the costs used for the no treatment group in the model, and how these were derived.

We reproduce Table 53 below for convenience.

Table 16: Intervention and comparator costs (reproduced from Table 53 of CS)

Drug	Cost per day	Annual cost
Daridorexant 50mg	[redacted]	[redacted]
Daridorexant 50mg*	NA	[redacted]
No Treatment	£0	£0

*Discontinuation-adjusted 12-month cost
NA=not applicable

a) **Response:** Treatment with daridorexant costs [redacted] per day x 365 days per year giving an annual treatment cost of [redacted]. In our response to B6 we justify why the only appropriate comparator for daridorexant is no treatment which has an intervention cost of £0.

b) **Response:** Discontinuation adjustments are clarified in detail in our response to non-priority question B10 above, in particular the added table in part (b).

c) **Response:** In common with standard modelling practice, background disease-related costs are included in the no treatment arm of the model. Since these are disease-related and not intervention-related, these were not included in Table 16 which was a NICE template table asking specifically for 'intervention' costs. The calculation of these disease-related costs is covered in our response to non-priority question B16 below.

B16. CS section B.3.5.2 *"Health state unit costs and resource use"* does not provide the health states that were considered within the economic model. This section also provides unit costs for healthcare resource use for 2021 (CS Table 54) but does not provide details regarding healthcare resource utilisation or price inflation adjustments.

a) Please specify the different health states that were considered and details regarding the costs within each of these health states.

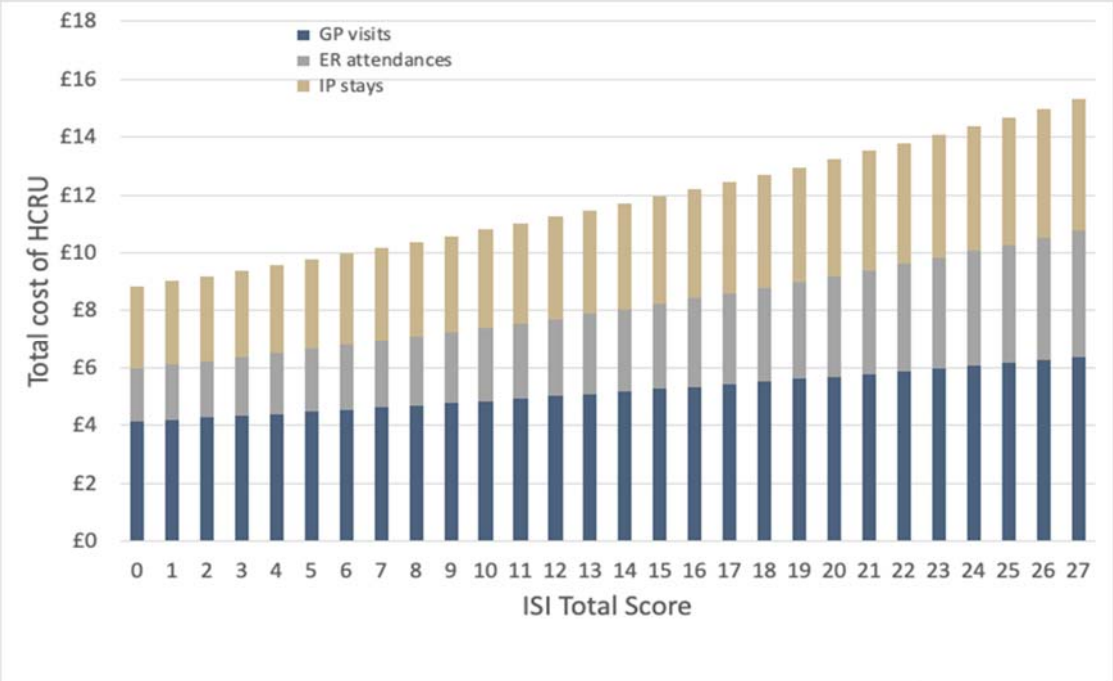
b) Unit costs for 2021 for emergency room and inpatient care are derived from NHS England 2019/20 costs. Please provide detail as to how these costs were adjusted for inflation.

c) CS Figure 3 is referred to for the predicted direct healthcare cost of each category of resource use. Please provide detail regarding how resource utilisation within each given category was estimated.

d) Please provide justification for the three given resource categories (CS Table 54) being the only costs (in addition to drug acquisition costs) relevant to the NHS and Personal Social Services perspective.

a) **Response:** As described in our response to priority question B1 above, the model presented does not have 'health states'. Instead, there is a direct mapping function between ISI© total score and health care resource use / cost and the cost for health resource use category is represented in Figure 3 (reproduced below for convenience).

Figure 9: Total predicted direct healthcare costs by ISI© score (49) (shown as the value of a one-point reduction to that score; reproduced from Figure 3 of CS)



ER = Emergency room; GP = General practitioners; IP = Inpatient; ISI = Insomnia Severity Index.

b) **Response:** Costs were inflated using the CPI index 06: Health (50).

c) **Response:** Our apologies – this was not adequately described in the submission – although all of the calculations are provided within the model on worksheet <*Ref* | *NHWS HCRU*>. Resource use information was estimated in the same way for each of the three categories of resource use. Resource use in terms of number of visits was the explanatory variable, with ISI total score as an explanatory variable and a range of other covariates included in the model as potential confounders in order to estimate the independent effect of ISI on resource use. The model was a GLM with negative binomial distribution family (commonly employed for count data with overdispersion) and a log link (which is the canonical link for the distribution). The estimated coefficients of the

models are presented in column C with the corresponding standard errors in column D. The covariance matrices and the Cholesky decomposition matrix for the models are presented alongside the coefficient estimates on the same worksheet.

At the top of the worksheet, in columns H, I and J you find the simplified models of resource use where the linear predictor is collapsed using the average values of all the potential confounding variables to give a 'new intercept' variable along with the independent coefficient for the impact of ISI total score on the resource use category. These are used to calculate the figures, by ISI© score to populate figure 3 above, but these are also passed to the *<Model | Parameters>* worksheet to drive the disease cost estimation in the model.

d) **Response:** These are the only cost categories that were available in the NHWS data. While other costs (e.g. concomitant medication use) are relevant to the NHS and PSS perspective, these costs were not available in the NHWS data set. However, since the disease costs are background costs that improved ISI© will offset, all missing cost categories reflect a conservative assumption that acts against treatment.

B17. The CS includes a scenario analysis which incorporates costs associated with productivity losses into the economic model. Productivity losses from chronic insomnia disorder were estimated in two alternative ways.

a) Please provide a step-by-step explanation as to how the productivity losses were calculated.

b) For the estimation derived directly from the Sheehan Disability Scale (SDS) included in the clinical programme, please provide justification for the assumption of 4.9 working days per week applied to the level of absenteeism.

c) Table 55 of the CS contains productivity losses estimated from the SDS. Please provide justification for the baseline productivity losses being higher for the treatment group, compared with the placebo group.

d) Please provide an additional scenario analysis using a more conservative estimate of the average for working days per week.

Section B.3.5.4 of the submission details the two ways in which the productivity losses are calculated in the model. We welcome the opportunity to provide more detail on their calculation.

a) **Response:** (1) Sheehan Disability Scale (SDS)

A five-item index that asks three questions about how work/school life has been affected, how social time has been affected and how family time has been affected and then asks two questions about the number of days lost and the number of days underproductive. Days lost is labelled absenteeism. The work/school life level of underproductivity multiplied by the number of days underproductive is labelled presenteeism. Together they combine to give a total number of days lost. The timescale is for the week prior to administration of the instrument.

For the SDS© results, please note that Tables 25 and 26 of the CS, the number of lost days, the number of underproductive days and the level of underproductivity at work/school reported in the previous week for the phase III 301 study. For the 303 extension study, the equivalent tables for SDS are tables 45 and 46 of the CS. Tables 24 and 44 of the CS show the productivity score. These three items are used to calculate the productivity losses.

These results are combined on worksheet <Ref | SDS> in the model with the observed responses for each of the three items that are used to calculate productivity losses at each time point in each arm of the model and the differences between the two arms.

The first step is to combine the days underproductive with the level of underproductivity to create a whole time equivalent (WTE) number of productive days lost. The ten-point underproductivity scale (0 completely underproductive, 10 completely productive) is divided by 10 and multiplied by the days underproductive to give the WTE losses due to presenteeism.

The second step is to add the days of absenteeism to give total days lost due to both absenteeism and presenteeism in the previous week.

The third step is to apply a unit cost to translate these days into productivity losses measured in GBP. To do this we assume that there are 255 working days in the year, and we divide this by 52 weeks to get 4.9 working days per week. We divide the working days per week by 7 to get the proportion of the week spent working and multiply this by first the number of weeks in the time period and then the days lost to get the working days lost in the time period. We then multiply the working days lost by the estimated 7.5 hours in a working day and apply the hourly wage rate to get the productivity losses in the time period. We do this separately by treatment arm and for absenteeism and presenteeism before calculating the total cost per time period and summing up over time period to get the productivity losses over the year.

b) **Response:** (2) Work Productivity and Activity Impairment (WPAI) questionnaire. Productivity losses data from the NHWS were collected using the Work Productivity and Activity Impairment (WPAI) questionnaire. The WPAI-GH consists of six questions: 1 = currently employed; 2 = hours missed due to health problems; 3 = hours missed for other reasons; 4 = hours actually worked; 5 = degree health affected productivity while working (using a 0 to 10 Visual Analogue Scale (VAS)); 6 = degree health affected productivity in regular unpaid activities (VAS). The recall period for the questions 2 to 6 is seven days. Absenteeism and presenteeism outcomes were generated from the WPAI-GH and expressed in percentages by multiplying the following scores by 100, using the following respective equations: 1) percent work time missed due to health (absenteeism) = Q2/(Q2 + Q4) for those who were currently employed; 2)

percent impairment while working due to health (presenteeism) = Q5/10 for those who were currently employed and actually worked in the past seven days.

Lost productivity costs were calculated using the following equations:

Cost of absenteeism = hourly wage * hours missed work in past 7 days in NHWS * work weeks per year

Cost of presenteeism = hourly wage * (lost productivity while at work in NHWS/10) * hours worked in past 7 days in NHWS * work weeks per year

Total indirect cost = sum of cost of absenteeism and cost of presenteeism

c) **Response:** The median annual wage for 2021 is £25,971. Like for the SDS, given 255 working days per year in the UK and 7.5 hours per working day in the UK, the median hourly earnings in 2021 for all employees is £13.58 per hour.

d) **Response:** With 255 working days in the year and 52 weeks in the year we get 4.9 working days per week.

As with other variables measured at baseline there can be slight imbalances. Overall, the difference is small and non-significant and amounts to just £4.41 of difference between the two arms at baseline.

We presume the EAG is objecting to the 255 working days per year assumption. Perhaps because some of those working days can be taken by employees as vacation. So assuming 30 days of vacation time per year we could assume 225 / 52 = 4.3 working days per week. However, to get the hourly wage rate we divide the median UK annual salary by 255 working days and assume 7.5 hours per day. (See cell D86 on the <Model | Parameters> worksheet.). Since any change to the assumed working days per year would change in both places our estimates of productivity loss do not change.

Scenario and sensitivity analyses

B18. To address uncertainty surrounding the cost effectiveness results the company performed deterministic and probabilistic sensitivity analyses, as well as scenario analyses. However, for the deterministic sensitivity analysis and scenario analyses, it is unclear which input parameters and/or assumptions were changed/varied and which values were used.

Please provide an overview of the specific parameters (and structural assumptions) that were changed/varied and their corresponding values.

Response: We agree that the deterministic sensitivity analysis could be confusing because of the type of model so we welcome the opportunity to clarify. Please note that Table 56 of the CS gives the full parameter list and we assume it is clear that all parameters assigned a distribution in Table 56 were included in the probabilistic analysis. The deterministic sensitivity analysis is presented as a Tornado Diagram in Figure 20 of the CS. We are pleased to confirm that the label 'ISI values 301' relates to letting all of the baseline, month 1 and month 3 ISI parameters that are listed in Table 56 and which are estimated from the phase III 301 study vary simultaneously in a probabilistic analysis while holding all other parameter values constant at their point estimates. Similarly, the label 'ISI values 303' in Figure 20 relates to letting all of the ISI parameters that are listed in Table 56 and which are estimated from the 303 extension study vary simultaneously in a probabilistic analysis while holding all other parameter vales constant at their point estimates. We do realise, as we state in the CS, that this is not a usual deterministic analysis. The principle is that we are letting all of the parameters that are jointly estimated/ related vary together while holding all other parameters constant. We hope that this clarifies the situation. We are not varying the structural assumptions of the model though in Figure 20. Structural changes to the model are presented as scenarios. Although the full parameter listing for scenarios is not listed in Table 56, it should be very clear from the model itself. For example, the sub-group

analysis uses the interacted SUR model that is presented in full (with Cholesky decomposition matrix) in worksheet <*Ref* | *ISI analysis (severity)*>.

Validation and transparency

B19. Several regression models are used as key structures in the model. Relatively little detail is provided about their conduct and the adherence to good practice guidelines.

a) Please list all regression models described in the CS.

b) Please complete the Checklist for Statistical Regression Analyses proposed by Kearns et al. 2013 (Appendix in https://doi.org/10.1007/s40273-013-0069-y) for each regression model listed in the preceding question.

c) Please elaborate on any shortcomings in the reporting of the regression analyses that were identified by the Checklist and provide additional clarifications and/or justification when applicable.

a) **Response:** The following table lists all regression models employed in the CS

Table 17: Lists of regression models employed in the CS

No.	Description	N
1.	Seemingly unrelated regression of study 301	557
2.	Mapping algorithm from ISI to EQ5D utility	17,955
3.	Mapping function from ISI to GP visits	2,306
4.	Mapping function from ISI to ER visits	2,306
5.	Mapping function from ISI to IP visits	2,306

6.	Mapping function from ISI to WPAI (absenteeism)	1,106
7.	Mapping function from ISI to WPAI (presenteeism)	1,106

b) **Response:** The Kearns checklist has already been reported for the first equation in the response to question B8 above. The remaining regression models were all estimated by a third-party vendor using a common approach to the modelling. The Kearns checklist is completed below to the best of our ability (as end user of the provided models) and applies to all of regression models 2 through 6.

Application of the Kearn's checklist to regression models used in the NHWS data set.

1. Have the objectives of the analysis been stated?

Yes. The objective was to provide a mapping between ISI and the dependent variable of interest while adjusting for all possible observed potential confounders.

2. Has the need for a de novo regression analysis been justified?

Yes. The regressions were necessary because HRQoL utility and HCRU was not measured in the clinical trial programme. The possible exception in regression 6 as SDS productivity was measured in the trial – however, there was an opportunity to examine a further approach to estimating productivity by mapping through to WPAI.

3. Has the source of the data used been stated? This would include synopses of key study features such as socio-demographic/clinical characteristics and the data collection method.

Yes. The NHWS data set is described in the CS.

4. Has the total sample size available been reported?

Yes. See summary table in (a) above.

5. Are sufficient explanations of all variables used provided?

Yes. The key explanatory variable is ISI – other explanatory variables available in the data were treated as potential confounders and used to ensure the mapping between ISI total score and the dependent variable of interest was independent of observed confounders.

6. Are sufficient numerical and/or graphical summaries provided?

Yes. Detailed numerical reporting in the Excel model (including covariate matrices and Cholesky decomposition matrices) with graphical summaries in the CS.

7. Has the quality of data (missing values, outliers, possible bias, etc.) been described?

Yes, it has been described in the NHWS report. The NHWS is the largest international self-reported patient database in the healthcare industry, with annual survey responses dating back to 1998 in the US, 2000 in Europe and 2008 in Asia (51). The protocol and questionnaire for the NHWS have been reviewed and have been granted exemption status by Pearl Institutional Review Board (IRB; Indianapolis, IN, USA; 19-KANT-204).

NHWS does not include “true” missing values, as values are missing due to survey skip logic. For example, respondents who reported not taking prescription medication to treat insomnia would not have response data for questions regarding prescription medication use. Participant responses were collected via computerized interface, which eliminated the likelihood of missing data because an item must be answered before the next one is displayed. Based on the programming of the survey, out-of-range or implausible responses are not possible. Prior to initiating the survey, appropriate edit programming was conducted to assure the final dataset requires minimal cleaning of invalid responses. The questionnaire was designed so that instructions are as easy to understand and clear as possible to avoid missing data. These

programming procedures for the web-based survey data entry tool included response ranges, consistency checks, skip patterns, and other special edit procedures where applicable. At each step of data processing, results or data manipulations were cross checked by Cerner Enviza team members who independently replicated the results and/or verified that the data have been handled appropriately and accurately. Any inconsistencies identified during this process were corrected before any further analysis was completed.

Complete data were available for all items except those allowing “don’t know” or “refuse to answer” responses (e.g., sensitive questions), if applicable. In such cases, if those variables were included as outcome measures in bivariate analyses or as covariates in multivariable models, missing values were included as a separate, defined category, or assimilated into another category, or omitted altogether (depending on whether either approach was necessary, e.g., due to problems with model convergence). If those variables were analysed as outcomes, respondents with missing data were excluded from analysis (and the subsample for analysis were reported). No methods to impute missing values were applied.

8. Has the type/method of regression model(s) considered been stated/justified?

Yes. Generalised linear models with appropriate distribution families and canonical link functions.

9. Have any modelling assumptions been stated?

Yes. Functional form of the model and choice of covariates.

10. Is a convincing rationale given for the inclusion of explanatory variables?

Yes. The key explanatory variable is ISI – other explanatory variables available in the data were treated as potential confounders and used to

ensure the mapping between ISI total score and the dependent variable of interest was independent of observed

11. Are sufficient details about the computational methods used provided?

Yes, models were estimated in SAS using the `genmod` procedure.

12. If more than one model was considered, has justification been given for why the preferred model has been selected?

N/A as only one model per regression was considered.

13. Has the choice of covariates been justified?

Yes. The key explanatory variable is ISI – other explanatory variables available in the data were treated as potential confounders and used to ensure the mapping between ISI total score and the dependent variable of interest was independent of observed confounders. analysis.

14. Is the sample size reported for every model presented?

Yes.

15. Has the handling of missing values (if any) been described?

As mentioned above in answer 7., no methods to impute missing values were applied, because the NHWS does not include “true” missing values, as values are missing due to survey skip logic. Complete data were available for all items except those allowing “don’t know” or “refuse to answer” responses (e.g., sensitive questions), if applicable. In such cases, if those variables were included as outcome measures in bivariate analyses or as covariates in multivariable models, missing values were included as a separate, defined category, or assimilated into another category, or omitted altogether (depending on whether either approach was necessary, e.g., due to problems with model convergence). If those variables were analysed as outcomes,

respondents with missing data were excluded from analysis (and the subsample for analysis were reported).

16. Are the coefficient estimates provided?

Yes. With full reproduction in the Excel model.

17. Are appropriate measures of uncertainty and significance provided?

Yes.

18. Are summary measures of goodness of fit presented?

No. Goodness of fit not formally assessed.

19. Are details of the results of a residual analysis provided?

No.

20. Has the model been validated on external (or quasiexternal) data?

No. Not clear that any external data were available.

21. Is the plausibility of the modelled predictions and/or coefficients discussed?

Yes. The estimates are fully utilised and presented and their plausibility is the subject of the modelling and associated CS.

22. Are the results compared to the literature and/or other data?

To a limited extent. The utility mapping is not disimilar to the previously reported Gu et al paper and the WPAI estimates are compared to the observed SDS estimates from the trial and discussed.

23. Has the method for handling parameter uncertainty been reported?

Yes. Covariance matrices / Cholesky decomposition matrices are used to propagate uncertainty in the probabilistic analysis of the decision model.

24. Is sufficient detail given for how parameter uncertainty was handled (e.g. if a variance–covariance matrix is used, is this available in some form?)

Yes. The covariance matrices are included in the Excel model and are used to calculate the Cholesky decomposition matrices to propagate the uncertainty.

25. Is parameter uncertainty appropriately reflected in the DAM?

Yes, the regression model is used to get the appropriate statistical quantities to propagate parameter uncertainty.

26. Has any structural (model) uncertainty been explored (in the DAM)?

No. Only single regression models were fit to the data. The exception to this is the utility algorithm where there was more than one regression model fit but only the preferred model utilised.

27. Have the model's limitations been discussed (and explored if possible)?

Only to a very limited extent.

c) **Response:** When evaluated against the Kearns checklist there are some minor shortcomings of the regression modelling identified. These are principally that only a single functional form model was employed (with the exception of the utility regression) and there could have been more extensive reporting of the model fit statistics. Nevertheless, the general conclusion should be that these regression equations are fit for purpose. They describe the relationship between ISI© and the dependent variable of interest based on a large cross-sectional survey and each model has a functional form that is appropriate to the dependent variable

and all available covariates in the dataset are used to protect from confounding of the relationship between the key covariate of interest (ISI© total score) and the dependent variable. Of course, as this is an observational dataset the risk of unobserved confounding remains, but the main risk of confounding by selection is not relevant to this aspect of the modelling.

B20. In section 3.14 of the CS, it is described that the face validity and technical validity of the economic model was assessed.

a) Please provide a detailed description of the validity assessment performed as well as the results.

b) Please complete the TECH-VER checklist (Büyükkaramikli et al. 2019, https://doi.org/10.1007/s40273-019-00844-y) and provide the results.

Response: The validation process followed and described in the CS did not make direct reference to the TECH-VER guidance as described by Büyükkaramikli et al (52). However, we now restate the validation process below with reference to the TECH-VER guidance in Table 1 of the paper.

Pre-analysis assessment of completeness and consistency was done first, including checking for hidden sheets, ranges, and modules, links to external programs, hardcoded values, and password protection. No changes were made from this review. We then assessed calculation consistency between the model and the description and values reported in or derived from the trial data driving the model. Again, no inconsistency was found.

The model was then assessed for correctness of implementation. Black-box testing was used to identify any suspicious values, and white-box testing was used to further investigate formulas as necessary. Two errors were identified and remedied at this stage (note these are listed on the version control section of the title worksheet of the model – version 2.22 was the version checked and the corrected version was 2.23):

1. ICER net productivity values miscalculated: The ICER net productivity values at each time point were calculated as 12-month values instead of reflecting the time between each data measure.

2. Net health benefit miscalculated: The net health benefit calculations in the model were calculating relative to a 3-month time period instead of the appropriate 12- month time frame.

Next, event calculations were checked to ensure that the distribution of cohorts among ISI© subgroups and severity subgroups reflected the clinical trial results, and that the appropriate costs and QALYs were applied to relevant severity subgroups. No errors were found.

The results calculations were then assessed. This confirmed that the summation of costs, life years, and quality-adjusted life years were functioning correctly and pulling in the correct data from the model, and that the ICER calculations were correct. We also confirmed that cost and quality of life discounting were appropriately addressed in the lifetime model scenario.

Finally, the model uncertainty analysis and scenario analyses were validated. No errors were found in the functioning of the macro utilised to run the probabilistic sensitivity analysis. All scenario analysis toggles and macros were also found to be sufficiently functional, as was all interactive model functionality.

B21. Please provide cross validations, i.e. comparisons with other relevant NICE TAs focussed on similar, potentially relevant, diseases, e.g. related NICE recommendations and NICE Pathways listed in the final scope, and elaborate on the identified differences regarding:

a) Model structure and assumptions, input parameters related to clinical effectiveness, health state utility values, resource use and costs

b) And how these differences affect estimated outcomes per comparator / interventions (life years, QALYs, costs)

Response: As described in our response to non-priority question B14 above, the models presented in TA77 and MTG70 are not full cost-per-QALY models for insomnia and therefore our results are not directly comparable to the previous NICE guidance in this area.

B22. Further external validation of modelled effectiveness would be desirable.

a) Please report on the face validity of the model structure, model assumptions, model inputs, intermediate outcomes as well as final outcomes in more detail (including what aspects were assessed and what were the considerations as well as conclusions).

b) Please assess the external validity of model inputs, intermediate outcomes as well as final outcomes using

1) Evidence used to develop the economic model.

2) Evidence not used to develop the economic model.

a) **Response:** As described in the CS, face validity was tested through an Advisory Board including two expert clinicians and three experienced health economics experts. Ultimately, the face validity of the model depends on two factors: the validity of ISI© as a measure of insomnia severity (we believe it is as it is perhaps the most widely used insomnia index having been translated into over 50 languages); and the validity of the mapping function from ISI© to EQ5D (again we believe we have a function that follows best practice guidance and is consistent with the NICE methods guide).

b) **Response:**

(1) In terms of evidence used to develop the model: our short-term model uses direct clinical trial evidence on the effect of treatment on ISI from the phase III 301 study and the 303 extension study. RCTs represent a high standard of evidence and the design was accepted by the regulator. The mapping function

was developed using a unique commercial data set that administered the ISI instrument alongside the EQ5D instrument.

(2) In terms of evidence not used to develop the model we can point to the fact that ISI is an accepted measure for insomnia severity. It has been used in many clinical studies over many years and has undergone extensive validation over that time. The only external evidence that we are aware of on the relationship between ISI© and EQ5D is the Gu mapping function (47). At the EAG's request we have updated version 2.30 of the CEM to include the functionality to use the Gu algorithm as an alternative to the NHWS algorithm we developed and employed in the model. Using the Gu algorithm results in a very small increase in the estimated ICER (from [redacted] to [redacted]).

Severity

B23. In CS section B.3.6 Severity, it is stated *"Not relevant to this submission"*

a) Please justify this statement.

b) Please elaborate on the disease severity and estimate the absolute QALY shortfall (AS) and proportional QALY shortfall (PS), providing calculation details ensuring reproducibility of the AS and PS.

c) Please justify what severity modifier is appropriate.

Response: The QALY shortfall does not justify an additional severity weighting and we therefore assume a QALY weight of 1 will apply.

Proportional Shortfall (PS)

EQ5D quality of life norms (Schneider et al, 2022 Table 1) for women and men are 0.802 and 0.837 respectively at age 55 (mean age in the 301/303 studies). This gives an approximate 0.819 Quality of life norm at age 55 averaged across both sexes.

In the model (based on the observed data for study 301 at end of follow-up) the mean ISI in the placebo group is 13.9 corresponding to an EQ5D utility of 0.698.

PS = 1 – 0.698/0.819 = 15% which is an order of magnitude below the 85% threshold for qualifying for a PS multiplier. (Even if we used baseline ISI of 19.2 corresponding to an EQ5D utility of 0.613 the shortfall is only 25%).

Absolute Shortfall (AS)

Table 2 of the Schneider et al paper gives the quality-adjusted and discounted life expectancy for women and men at age 55 as 14.19 and 13.77 respectively or 13.98 averaged across both sexes. Applying the 15% and 25% PS figures to these absolutes gives the absolute shortfalls as 2.1 and 3.5 absolute QALY shortfalls which are both far below the 12 QALY shortfall required for the multiplier to apply.

Uncertainty

B24. In CS section B.3.7 Uncertainty, it is stated “Not relevant to this submission”.

a) Please justify this statement.

b) Please elaborate on the uncertainty, both quantified and non-quantified uncertainty, surrounding the cost effectiveness estimates.

c) Please elaborate on potential biases that could potentially increase the estimated ICER.

Response: We had understood that this section was for new technologies where there may be significant gaps in the evidence base (e.g. single arm trials) such that the level of uncertainty merits entry into a managed access scheme. This is not the case with daridorexant where there is strong clinical trial data that shows a treatment impact on insomnia and a strong mapping mechanism that suggests insomnia is associated with EQ-5D utility. We have discussed uncertainty in each of the scenarios presented in the CS. As the base case is a 12-month model, by design it mitigates against uncertainty. This was

discussed in detail in the decision problem meeting - it was agreed that this was a specific benefit of our approach and that previous appraisals highlighted the uncertainty of long-term modelling in insomnia.

References

1. Wilson S, Anderson K, Baldwin D, Dijk DJ, Espie A, Espie C, et al. British Association for Psychopharmacology consensus statement on evidence-based treatment of insomnia, parasomnias and circadian rhythm disorders: An update. J Psychopharmacol. 2019;33(8):923-47.
2. Substance Abuse and Mental Health Services Administration. Table 3.36, DSM-IV to DSM-5 Insomnia Disorder Comparison. CBHSQ Methodology Report. Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration Substance Abuse and Mental Health Services Administration, Rockville, MD.; 2016
3. Sateia MJ. International Classification of Sleep Disorders-Third Edition. CHEST. 2014;146(5):1387-94.
4. Ree MJ, Richardson C. Insomnia disorder update: the benefits of screening and treatment for this common presentation. Internal Medicine Journal. 2021;51(11):1798-805.
5. Buysse DJ, Thompson W, Scott J, Franzen PL, Germain A, Hall M, et al. Daytime symptoms in primary insomnia: a prospective analysis using ecological momentary assessment. Sleep Med. 2007;8(3):198-208.
6. Hays RD MR, Spritzer K, Drake C, Roth T. Psychometric properties of the Restorative Sleep Questionnaire and Daytime Consequences of Sleep Questionnaire. Associated Professional Sleep Societies (APSS-19th Annual Meeting); Denver (CO)2005.
7. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep. 1997;20(10):835-43.
8. Moul D, Pilkonis P, Miewald J, Carey T, Buysse D. Preliminary study of the test-retest reliability and concurrent validities of the Pittsburgh Insomnia Rating Scale (PIRS). Sleep. 2002;25:A246-A7.
9. Yeun EJ, Shin-Park KK. Verification of the profile of mood states-brief: cross-cultural analysis. J Clin Psychol. 2006;62(9):1173-80.
10. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2001;2(4):297-307.
11. Hudgens S, Phillips-Beyer A, Newton L, Seboek Kinter D, Benes H. Development and Validation of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ). Patient. 2021;14(2):249-68.
12. Kendzerska TB, Smith PM, Brignardello-Petersen R, Leung RS, Tomlinson GA. Evaluation of the measurement properties of the Epworth sleepiness scale: a systematic review. Sleep Med Rev. 2014;18(4):321-31.
13. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep. 2011;34(5):601-8.
14. Data on file. UK Insights Deep Dive. 2022.
15. Data on file. IQVIA Longitudinal patient data (LPD) 2021.

16. Idorsia Pharmaceuticals Ltd. Data on file. Clinical Study Report. ID-078A301: Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder. 15 October 2020.
17. Sleepio to treat insomnia and insomnia symptoms. Medical technologies guidance [MTG70]: NICE; 2022 Available from: https://www.nice.org.uk/guidance/mtg70.
18. Rob Newson AD. Cognitive Behavioral Therapy for Insomnia (CBT-I): Sleep Foundation; 2022 Available from: https://www.sleepfoundation.org/insomnia/treatment/cognitive-behavioral-therapy-insomnia.
19. Matthews EE, Arnedt JT, McCarthy MS, Cuddihy LJ, Aloia MS. Adherence to cognitive behavioral therapy for insomnia: a systematic review. Sleep Med Rev. 2013;17(6):453-64.
20. Heesch CB. The long-term use of sedative hypnotics in chronic insomnia. Mental Health Clinician. 2014;4(2):78-81.
21. Sleepio to treat insomnia and insomnia symptoms. Medical technologies guidance [MTG70]: NICE; 2022 Available from: https://www.nice.org.uk/guidance/mtg70.
22. Morgenthaler T, Kramer M, Alessi C, Friedman L, Boehlecke B, Brown T, et al. Practice parameters for the psychological and behavioral treatment of insomnia: an update. An american academy of sleep medicine report. Sleep. 2006;29(11):1415-9.
23. Dündar Y, Dodd S, Strobl J, Boland A, Dickson R, Walley T. Comparative efficacy of newer hypnotic drugs for the short-term management of insomnia: a systematic review and meta-analysis. Human Psychopharmacology: Clinical and Experimental. 2004;19(5):305-22.
24. Holbrook AM, Crowther R, Lotter A, Cheng C, King D. Meta-analysis of benzodiazepine use in the treatment of insomnia. Cmaj. 2000;162(2):225-33.
25. Glass J, Lanctôt KL, Herrmann N, Sproule BA, Busto UE. Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits. BMJ. 2005;331(7526):1169.
26. Chen PL, Lee WJ, Sun WZ, Oyang YJ, Fuh JL. Risk of Dementia in Patients with Insomnia and Long-term Use of Hypnotics: A Population-based Retrospective Cohort Study. PLOS ONE. 2012;7(11):e49113.
27. Louzada LL, Machado FV, Nóbrega OT, Camargos EF. Zopiclone to treat insomnia in older adults: A systematic review. Eur Neuropsychopharm. 2021;50:75-92.
28. Ferracioli-Oda E, Qawasmi A, Bloch MH. Meta-analysis: melatonin for the treatment of primary sleep disorders. PLoS One. 2013;8(5):e63773.
29. Wade AG, Ford I, Crawford G, McMahon AD, Nir T, Laudon M, et al. Efficacy of prolonged release melatonin in insomnia patients aged 55–80 years: quality of sleep and next-day alertness outcomes. Current Medical Research and Opinion. 2007;23(10):2597-605.
30. Mignot E, Mayleben D, Fietze I, Leger D, Zammit G, Bassetti CLA, et al. Safety and efficacy of daridorexant in patients with insomnia disorder: results from two multicentre, randomised, double-blind, placebo-controlled, phase 3 trials. Lancet Neurol. 2022;21(2):125-39.
31. Idorsia Pharmaceuticals Ltd. Data on file. Final Clinical Study Report. ID-078A303: Multi-center, double-blind, parallel-group, randomized, placebo-controlled,

three doses, 40-week extension to studies ID-078A301 and ID-078A302 to assess the long term safety and tolerability of ACT-541468 in adult and elderly subjects with insomnia disorder. 2021.
32. Summary of Product Characteristics. QUVIVIQ.
33. Idorsia Pharmaceuticals Ltd. Data on file. Clinical Study Report. ID-078A301. Daridorexant number needed to treat (NNT) per objective and subjective sleep measure at month 1 and 3 in daridorexant 25 mg and daridorexant 50 mg 2022.
34. Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep. 2013;36(11):1747-55.
35. Withrow D, Roth T, Koshorek G, Roehrs T. Relation between ambulatory actigraphy and laboratory polysomnography in insomnia practice and research. J Sleep Res. 2019;28(4):e12854-e.
36. Idorsia Pharmaceuticals Ltd. Data on File. AC-078A201: Multi-center, double-blind, randomized, placebo-controlled, active-reference, parallel-group, polysomnography dose-response study to assess the efficacy and safety of ACT-541468 in adult subjects with insomnia disorder. 2018.
37. Kuntz KM, Kitch BT, Fuhlbrigge AL, Paltiel AD, Weiss ST. A novel approach to defining the relationship between lung function and symptom status in asthma. J Clin Epidemiol. 2002;55(1):11-8.
38. Briggs AH, Bousquet J, Wallace MV, Busse WW, Clark TJH, Pedersen SE, et al. Cost-effectiveness of asthma control: an economic appraisal of the GOAL study. Allergy. 2006;61(5):531-6.
39. Briggs AH, Baker T, Risebrough NA, Chambers M, Gonzalez-McQuire S, Ismaila AS, et al. Development of the Galaxy Chronic Obstructive Pulmonary Disease (COPD) Model Using Data from ECLIPSE: Internal Validation of a Linked-Equations Cohort Model. Medical Decision Making. 2016;37(4):469-80.
40. Yin J, Jin X, Shan Z, Li S, Huang H, Li P, et al. Relationship of Sleep Duration With All-Cause Mortality and Cardiovascular Events: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. Journal of the American Heart Association. 2017;6(9):e005947.
41. Draft scope and provisional stakeholder list comments table (post-referral). Single Technology Appraisal (STA). Daridorexant for treating insomnia disorder: National Institute for Health and Care Excellence 2021 Available from: https://www.nice.org.uk/guidance/gid-ta10888/documents/draft-scope-and-provisional-matrix-comments-table-post-referral.
42. NICE Guideline [NG202]. Obstructive sleep apnoea/hypopnoea syndrome and obesity hypoventilation syndrome in over 16s: National Institute for Health and Care Excellence 2021 Available from: https://www.nice.org.uk/guidance/ng202.
43. McDaid C, Griffin S, Weatherly H, Durée K, van der Burgt M, van Hout S, et al. Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis. Health Technol Assess. 2009;13(4):iii-iv, xi-xiv, 1-119, 43-274.
44. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. Health Econ. 2004;13(5):461-75.
45. Spronk I, Polinder S, Bonsel GJ, Janssen MF, Haagsma JA. Adding a fatigue item to the EQ-5D-5L improves its psychometric performance in the general population. Journal of Patient-Reported Outcomes. 2022;6(1):1.

46. Perneger TV, Courvoisier DS. Exploration of health dimensions to be included in multi-attribute health-utility assessment. Int J Qual Health Care. 2011;23(1):52-9.
47. Gu NY, Botteman MF, Ji X, Bell CF, Carter JA, van Hout B. Mapping of the Insomnia Severity Index and other sleep measures to EuroQol EQ-5D health state utilities. Health Qual Life Outcomes. 2011;9:119.
48. Technology appraisal guidance [TA77]. Guidance on the use of zaleplon, zolpidem and zopiclone for the short-term management of insomnia: National Institute for Health and Care Excellence 2004.
49. Idorsia Pharmaceuticals Ltd. Data on File. BV_Insomnia-NHWS Results_UK_v1.0_23March2021. Kantar2021.
50. Data and analysis from CENSUS 2021. CPI INDEX 06 : HEALTH 2015=100: Office for National Statistics; 2021 Available from: https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/d7bz/mm23.
51. National Health and Wellness Survey: Health economics; Available from: https://www.healtheconomics.com/resource/national-health-and-wellness-survey.
52. Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. Pharmacoeconomics. 2019;37(11):1391-408.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Daridorexant for treating insomnia disorder [ID3774]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Vicki Beevers**

2. Name of organisation	The Sleep Charity
3. Job title or position	CEO and Founder
4a. Brief description of the organisation (including who funds it). How many members does it have?	The Sleep Charity is a national organisation registered with the Charity Commission. It does not have a membership, its purpose is to empower the nation with sleep support. It is funded through various income streams including; grants, contracts with CCGs, donations, trading. We interact with individuals via our website and have a social media presence. We also run a national helpline and are commissioned to provide sleep services for families.
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.]	10/09/2021 - Takeda UK Ltd (Shire) - £125 17/06/2021 - Gilead Sciences Ltd - £540 01/04/2021 -Takeda UK Ltd (Shire) - £150 Funding has been received from the above pharmaceutical companies in return for the charity providing information around behavioural sleep intervention, this is provided through written information and presentations.

If so, please state the name of the company, amount, and purpose of funding.	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Through a variety of means. We have steering groups for various projects, we evaluate all of our one to one support/training. Gather experiences is an integral part of our work, we have also worked with various academic organisations to ensure the information that we gain is independently evaluated and methods used appropriate.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	It is extremely isolating and impacts on every area of wellbeing; mental health, physical health, emotional wellbeing. Chronic sleep deprivation can impact on a person's ability to take part in everyday activities, increasing risk of trips/falls/road traffic accidents. The immune system can also become compromised, and it impacts on weight. There are close links with sleep issues and anxiety/depression. Quality of life can be seriously compromised, some individuals have had to give up work, impacting on income/housing/relationships. Carers experience chronic sleep deprivation and fatigue when caring for someone with the condition.

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	They report that there is little in the way of specialist support or help. Sleeping tablets are often prescribed but are viewed as more of a sticking plaster rather than a solution. They are often exhausted as well and need one to one support to make changes
8. Is there an unmet need for patients with this condition?	There is a huge unmet need in this area. People want to share their sleep issues with somebody who empathises and has expertise in the field. This involves taking a thorough assessment which cannot possibly take place in a standard GP appointment. Patients want to understand why they are having issues and what strategies would be most effective in tackling the problem. Some patients are offered sleeping tablets, others may be signposted onto an app such as Sleepstation or Sleepio. The apps can be effective depending on the individual and the problem, there has to be motivation to use it however which can be difficult when exhausted. This seems to be a postcode lottery and not all GPs in the areas are aware of the apps when patients can access them. The Sleep Charity provide a national sleep helpline that runs 5 times each week. Others are asked to keep sleep diaries which is often reported as being unhelpful as no interpretation of the data is made.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Not aware of anyone using this currently.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Not aware of anyone using this currently.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	N/A
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	**N/A**

Other issues	
13. Are there any other issues that you would like the committee to consider?	N/A

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission:

- Insomnia is a significant health inequality that impacts widely
- Sleep deprivation impacts on wellbeing; mentally/physically/emotionally
- There is an enormous cost to society by not addressing insomnia, effective treatment will be hugely cost saving
- There is little specialist support available in the UK
- Patients are often prescribed sleeping tablets with may not be appropriate, Cognitive Behavioural Therapy for Insomnia should be tried first.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Daridorexant for treating insomnia disorder (review of GID-TA10888) [ID3774]

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors

Mark Perry, Reviews Manager, KSR Ltd, United Kingdom (UK)
Bram Ramaekers, Health Economist, Maastricht UMC+, NL
Willem Witlox, Health Economist, Maastricht UMC+, NL
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Lisa Stirk, Information Specialist, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC+, NL
Bradley Sugden, Health Economist, Maastricht UMC+, NL
Andrea Fernandez Coves, Health Economist, Maastricht UMC+, NL
Teebah Abu-Zarah, Health Economist, Maastricht UMC+, NL
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC+, NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to Robert Wolff, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 20/09/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 2017 13/56/28.

Declared competing interests of the authors

None.

Acknowledgements

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Perry M, Ramaekers B, Witlox W, McDermott K, Stirk L, Otten T, Sugden B, Fernandez Coves A, Abu Zarah T, Posadzki P, Ahmadu C, Armstrong N, Joore M, Wolff R. Daridorexant for treating insomnia disorder (review of TA10888) [ID3774]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Mark Perry acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Kevin McDermott acted as a systematic reviewer, critiqued the clinical effectiveness methods and evidence and led the writing of the clinical evidence sections of the report. Charlotte Ahmadu and Pawel Posadzki acted as systematic reviewers and contributed to the writing of the report. Bram Ramaekers acted as health economic project lead, critiqued the company's economic evaluation and led the writing of the health economics sections of the report. Nigel Armstrong acted as Health Economist and contributed to the writing of the report. Willem Witlox, Thomas Otten, Bradley Sugden, Andrea-Fernandez Coves, Teebah Abu-Zarah and Manuela Joore acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Lisa Stirk critiqued the search methods in the submission and contributed to the writing of the report. Robert Wolff critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse event
AESI	Adverse events of special interest
AiC	Academic in confidence
ALDVMM	Adjusted Limited Dependent Variable Mixture Model
BMI	Body mass index
CADTH	Canadian Agency for Drugs and Technologies in Health
CBT	Cognitive behavioural therapy
CBT-I	Insomnia-related cognitive behavioural therapy
CEA	Cost effectiveness analysis
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CiC	Commercial in confidence
CL	Confidence limit
CLAD	Censored Least Absolute Deviations
CNS	Central nervous system
COVID-19	Coronavirus disease 2019
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CYP3A4	Cytochrome P450 3A4
DAR	Daridorexant
DB	Double blind
DCSQ	Daytime Consequences of Sleep Questionnaire
DISS	Daytime Insomnia Symptom Scale
DORA	Dual orexin receptor antagonist
DOX	Doxepin
DSA	Deterministic sensitivity analyses
DSM	Diagnostic and statistical manual of mental disorders
DSM-5	Diagnostic and statistical manual of mental disorders, 5th edition
DSU	Decision Support Unit
EAG	Evidence Assessment Group
ECG	Electrocardiogram
ECM	Established clinical management
EED	Economic Evaluations Database
EQ-5D	European Quality of Life-5 Dimensions
EOT	End of trial
ESS	Epworth Sleepiness Scale
ESZ	Eszopiclone
EUR	Erasmus University Rotterdam
FAS	Full analysis set
FDA	Food and Drug Administration
FE	Fixing errors
FOSQ	Functional Outcomes of Sleep Questionnaire
FV	Fixing violations
GP	General practitioner
HCRU	Healthcare resource utilisation
HRQoL	Health-related quality of life
HSUV	Health state utility value
HTA	Health Technology Assessment

ICD	International Classification of Diseases
ICER	Incremental cost effectiveness ratio
ICSD	International Classification of Sleep Disorders
ICTRP	International Clinical Trials Registry Platform
IDSIQ	Insomnia Daytime Symptoms and Impacts Questionnaire
iNHB	Incremental net health benefit
ISB	Independent Safety Board
ISI	Insomnia Severity Index
ISPOR	The Professional Society for Health Economics and Outcomes Research
ITC	Indirect treatment comparison
ITT	Intention to treat
KSR	Kleijnen Systematic Reviews Ltd
LEM	Lemborexant
LMZ	Lormetazepam
LPS	Latency to persistent sleep
LSEQ	Leeds Sleep Evaluation Questionnaire
LSO	Latency to sleep onset
LSM	Least squares mean
MD	Mean difference
MedDRA	Medical Dictionary for Regulatory Activities
MEL	Melatonin
MeSH	Medical subject headings
MJ	Matters of judgement
N (n)	Number
N/A	Not applicable
NCC	National Cost Collection
NHS	National Health Service
NHWS	National Health and Wellness Survey
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NNT	Number needed to treat
NR	Not reported
PBO	Placebo
PGA-S	Patient Global Assessment of Disease Severity
PGI-C	Patient Global Impression of Change
PICO	Population, intervention, comparator and outcome
PIRS	Pittsburgh Insomnia Rating Scale
POMS	Profile of Mood States
PRESS	Peer Review of Electronic Search Strategies
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analysis
PSG	Polysomnography
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PSQI	Pittsburgh Sleep Quality Index
PT	Preferred terms
PTSD	Post-traumatic stress disorder
QALY	Quality-adjusted life year
QoL	Quality of life

RAM	Ramelteon
RCT	Randomised controlled trial
REM	Rapid eye movement
RoB	Risk of bias
RR	Relative risk; Risk ratio
S1	Sleep stage 1
S2	Sleep stage 2
SAE	Serious adverse event
SD	Standard deviation
SDQ	Sleep Diary Questionnaire
SDS	Sheehan Disability Scale
SE	Standard error
SF-36	36-Item Short Form Survey
SFIS	Sleep Functional Impact Scale
SFRMS	Société Française de Recherche et Médecine du Sommeil
SLR	Systematic literature review
sLSO	Subjective latency to sleep onset
SmPC	Summary of product characteristics
SSRIs	Selective serotonin reuptake inhibitors
SSS	Stanford Sleepiness Scale
STA	Single Technology Appraisal
sTST	Subjective total sleep time
SUR	Seemingly unrelated regression
SUV	Suvorexant
sWASO	Subjective wake time after sleep onset
SWS	Slow wave sleep
TA	Technology Assessment
TEAE	Treatment emergent adverse events
TMZ	Temazepam
TRA	Trazodone
TSD	Technical Support Document
TST	Total sleep time
TZ	Triazolam
UK	United Kingdom
UMC	University Medical Centre
US / USA	United States of America
VAS	Visual Analogue Scale
WASO	Wake time after sleep onset
WHO	World Health Organization
WPAI	Work Productivity and Activity Impairment
WTE	Whole time equivalent
ZAL	Zaleplon
ZPC	Zopiclone
ZPD	Zolpidem

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Assessment Group (EAG) as being potentially important for decision making. If possible, it also includes the EAG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues, Section 1.2 presents the key model outcomes, Section 1.3 discusses the decision problem, Section 1.4 discusses issues relating to the clinical effectiveness, and Section 1.5 discusses issues relating to the cost effectiveness. A summary is presented in Section 1.6.

Background information on the technology, evidence and information on key as well as non-key issues are in the main EAG report. For more details, please see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (cost effectiveness).

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1.1: Summary of key issues

ID3774	Summary of issue	Report Sections
1	It is possible that the population in the trials is narrower than the population in the decision problem. This has implications for the applicability.	2.1
2	Although daridorexant is designed as a replacement treatment for those people that may be unsuitable for established care treatments such as CBT-I, most of those in the trials have never had the opportunity to receive or reject CBT-I.	2.1
3	The comparator in the decision problem is established clinical management. However, the comparator in the clinical effectiveness evidence presented in the CS is placebo with no mention of established clinical management and in the cost effectiveness section it is referred to as no treatment. There is no attempt by the company to perform an indirect treatment comparison to rectify this situation. The CS therefore fails to present data relating to the decision problem.	2.3
4	Numerous outcomes that measure the same construct are presented, increasing the risk of type I errors	2.4
5	The clinical effectiveness evidence (albeit evidence that covers daridorexant versus placebo rather than daridorexant versus established clinical management) omits a key paper	3.2
6	Ethnic make-up of the trials differs from the ethnic make-up of the UK population. The trials have not been sub-grouped for ethnicity sufficiently comprehensively across the two trials, making it difficult to exclude ethnicity as an effect modifier. Therefore, applicability of the trial findings is unclear.	3.2.1
7	Shorter term benefits of daridorexant over placebo do not appear to persist into the longer term in all cases	3.2.5
8	Studies 301 and 303 which inform the health economic model excluded patients with mental health problems. Because insomnia is frequently comorbid with other mental health problems the exclusion of patients with	4.2.3

ID3774	Summary of issue	Report Sections
	mental health problems may decrease the generalisability of the underlying evidence to the decision problem	
9	A variety of pharmaceuticals and therapies are available for the treatment of insomnia. The company only included no-treatment as a comparator to daridorexant in the health economic model.	4.2.4
10	The company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation.	4.2.4
11	As per the CS, the no-treatment arm was modelled to have no dropout, as patients receiving could not dropout from receiving no treatment. However, in the economic model provided by the company, the dropout rates observed in studies 301 and 303 for the daridorexant arm were applied to both daridorexant and no-treatment groups. This contradicts the statement made by the company.	4.2.6
12	For the company base case placebo effect was only included for the first three months in the no-treatment arm, but not for the remaining 40 weeks. The EAG considers that the effect of selective attrition on the daridorexant group and the possibility of regression to the mean on the no-treatment group, were not sufficiently justified by the company, and these effects could have biased the comparison in favour of the intervention.	4.2.7
13	The company excluded the AEs reported in studies 301 and 303 from their cost effectiveness model, assuming that these are minor AEs and would not be expected to have consequences on resource use or HRQoL.	4.2.7
14	There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.	4.2.8
15	In addition to treatment acquisition costs, the CS only incorporated costs and resource use for GP visits, emergency room visits and inpatient care. The company justified the decision due to these being the only categories captured in the NHWS dataset and stating that the approach was conservative. Such a conclusion cannot be drawn in the absence of supporting evidence.	4.2.9

AE = Adverse events; CBT-I = insomnia-related cognitive behavioural therapy; CS = company submission; EAG = Evidence Assessment Group; EQ-5D = European Quality of Life-5 Dimensions; GP = General Practitioner; HRQoL = Health-related quality of life; ISI = Insomnia Severity Index; NHWS = National Health and Wellness Survey; UK = United Kingdom

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life (QoL) in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost per QALY gained.

Overall, the technology is modelled to affect QALYs by:

- The insomnia severity index (ISI©) scores of Study 301 and Study 303,
- The ISI© score to European Quality of Life-5 Dimensions (EQ-5D) mapping algorithm.

Overall, the technology is modelled to affect costs by:

- Treatment costs
- Health care costs
- Productivity loss (in scenario analyses)

The parameters that have the greatest effect on the ICER (based on the company's sensitivity analyses) are:

- The ISI© scores of Study 301 and Study 303
- The parameters of the mapping algorithm of the ISI© scores to EQ-5D

Scenarios in the company submission (CS) that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- Inclusion of indirect costs (£[redacted] per QALY gained)
- Optimistic scenario (£[redacted] per QALY gained)
- Pessimistic scenario (£[redacted] per QALY gained)

1.3 The decision problem: summary of the EAG's key issues

The decision problem addressed in the CS is broadly in line with the final scope issued by NICE. However, the population is unclearly defined (Tables 1.2 and 1.3), the comparator in the trials differs from the NICE scope (Table 1.4) and there is a multiplicity of outcomes covering the same construct (Table 1.5).

Table 1.2: Key issue 1. Possibly inapplicable population

Report Section	**2.1**
Description of issue and why the EAG has identified it as important	It is possible that the population in the trials is narrower than the population in the decision problem. This has implications for applicability.
What alternative approach has the EAG suggested?	The company has been asked to confirm the definition of the population, to define the typical symptoms of insomnia and to define daytime impairment. The company has also been asked to discuss applicability implications if the population in the decision problem turns out to be broader than that defined in Study 301.
What is the expected effect on the cost effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	The company has been asked to confirm the definition of the population, to define the typical symptoms of insomnia and to define daytime impairment. The company has also been asked to discuss applicability implications if the population in the decision problem turns out to be broader than that defined in Study 301.
EAG = Evidence Assessment Group	

Table 1.3: Key issue 2. Study population not unsuitable for ECM such as CBT-I

Report Section	2.1
Description of issue and why the EAG has identified it as important	Although daridorexant is designed as a replacement treatment for those people that may be unsuitable for established care treatments such as CBT-I, most of those in the trials had never had the opportunity to receive or reject CBT-I.
What alternative approach has the EAG suggested?	A sub-group analysis would be useful that splits the sample into those using CBT-I and those that have not.
What is the expected effect on the cost effectiveness estimates?	This is currently uncertain.
What additional evidence or analyses might help to resolve this key issue?	A sub-group analysis would be useful that splits the sample into those using CBT-I and those that have not.
CBT-I = insomnia-related cognitive behavioural therapy; EAG = Evidence Assessment Group; ECM = established clinical management	

Table 1.4: Key issue 3. Inappropriate comparator

Report Section	2.3
Description of issue and why the EAG has identified it as important	The comparator in the decision problem is 'established clinical management'. However, the comparator in the clinical effectiveness evidence presented in the CS is placebo with no mention of established clinical management. It may be noted that concomitant treatments in the trials were allowed alongside the randomised treatments. CBT-I was allowed provided it had been started 4 or more weeks prior to baseline and continued throughout the studies. Non-prohibited drugs that were part of the patients' normal care were also permitted. There is little information provided on the comparability between arms. However, it can be assumed that because the studies were double-blinded any concomitant treatments should have been comparable between arms; blinding would ensure there could be no way in which preferential provision could be administered. Therefore, any ECM used in the studies would have been comparable between groups and so it could be regarded as a comparison of daridorexant plus ECM vs. ECM, which does not equate to daridorexant versus ECM (the comparison apparently defined in the NICE scope). In fact, the company argue in the cost-effectiveness section that the placebo arm is equivalent to no treatment. There is no attempt by the company to perform an indirect treatment comparison to rectify this problem. The CS therefore fails to present data relating to the decision problem. The first line clinical management for insomnia disorder is CBT-I, and it is unclear why this is not included in the decision problem as 'established clinical management'.
What alternative approach has the EAG suggested?	Unless the population precludes CBT-I, the company needs to carry out an indirect treatment comparison, using RCTs looking

Report Section	2.3
	at *CBT-I versus placebo plus no treatment or ECM excluding CBT-I* in a highly comparable population.
What is the expected effect on the cost effectiveness estimates?	The effects are currently uncertain.
What additional evidence or analyses might help to resolve this key issue?	Unless the population precludes CBT-I, the company needs to carry out an indirect treatment comparison, ideally anchored to RCTs looking at *CBT-I versus placebo* in a highly comparable population.
CBT-I = insomnia-related cognitive behavioural therapy; CS = company submission; EAG = Evidence Assessment Group; RCT = randomised controlled trial	

Table 1.5: Key issue 4. Multiple outcomes covering the same construct

Report Section	2.4
Description of issue and why the EAG has identified it as important	Numerous outcomes that measure the same construct are presented, increasing the risk of type I errors.
What alternative approach has the EAG suggested?	The company needs to select the most relevant of the multiple outcomes per construct (based on a reasoned rationale, not effect sizes)
What is the expected effect on the cost effectiveness estimates?	This is expected to reduce cost effectiveness estimates
What additional evidence or analyses might help to resolve this key issue?	The company needs to select the most relevant of the multiple outcomes per construct (based on a reasoned rationale, not effect sizes)
EAG = Evidence Assessment Group	

1.4 The clinical effectiveness evidence: summary of the EAG's key issues

The EAG identified three major concerns with the evidence presented on the clinical effectiveness, namely the omission of a key paper (Table 1.6), ethnic differences between the trials and the United Kingdom (UK) population (Table 1.7) and a possible lack of long-term benefit for some outcomes, which was not highlighted by the company (Table 1.8).

Table 1.6: Key issue 5. Omission of key paper

Report Section	3.2
Description of issue and why the EAG has identified it as important	The clinical effectiveness evidence (albeit evidence that covers daridorexant versus placebo rather than daridorexant versus established clinical management) omits a key paper. NCT02839200 (Dauvilliers et al. 2020) is included in the SLR, [Table 8 of Appendix D of the CS], but not in the main analysis of clinical efficacy evidence in the CS, even though it appears to be relevant, as it compares 50 mg daridorexant to placebo.
What alternative approach has the EAG suggested?	The key paper should be included and added to the evidence presented in the CS.
What is the expected effect on the cost effectiveness estimates?	The effects are uncertain.

Report Section	3.2
What additional evidence or analyses might help to resolve this key issue?	Inclusion of the key paper.
CS = company submission; EAG = Evidence Assessment Group; SLR = systematic literature review	

Table 1.7: Key issue 6. Ethnic differences between trials and UK population

Report Section	3.2.1
Description of issue and why the EAG has identified it as important	Ethnic make-up of the trials differs from the ethnic make-up of the UK population. The trials have not sub-grouped for ethnicity sufficiently comprehensively across the two trials, making it difficult to exclude ethnicity as an effect modifier. Therefore, applicability of the trial findings is unclear.
What alternative approach has the EAG suggested?	Comprehensive sub-grouping for ethnicity across both studies and all outcomes.
What is the expected effect on the cost effectiveness estimates?	The effect is uncertain.
What additional evidence or analyses might help to resolve this key issue?	Comprehensive sub-grouping for ethnicity across both studies and all outcomes.
EAG = Evidence Assessment Group; UK = United Kingdom	

Table 1.8: Key issue 7. Possible lack of long-term benefits for some outcomes that was not highlighted by the company

Report Section	Table 1.8, 3.2.5
Description of issue and why the EAG has identified it as important	Shorter term benefits of daridorexant over placebo do not appear to persist into the longer term in some outcomes. This was inadequately demonstrated in the CS. For example, the company carried out between-arm analyses for most of the shorter-term analyses (where significant effects were seen). However, for most of the longer-term analyses (where non-significant effects were subsequently demonstrated by between-arm analyses conducted by the EAG) the company failed to carry out between-arm analyses.
What alternative approach has the EAG suggested?	Although the lack of longer-term benefit could be related to a lack of statistical power in the longer-term analyses, this does not mean that a true lack of long-term benefit is excluded. The EAG has therefore stressed the importance of making the committee aware of the possibility of a lack of long-term benefit.
What is the expected effect on the cost effectiveness estimates?	This will reduce the cost effectiveness estimates.
What additional evidence or analyses might help to resolve this key issue?	Further longer-term data with greater statistical power would be very helpful.
CS = Company submission; EAG = Evidence Assessment Group	

1.5 The cost effectiveness evidence: summary of the EAG's key issues

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the EAG's summary and detailed critique in Section 4, and the EAG's amendments to the company's model and results are presented in Section 6. The key issues in the cost effectiveness evidence are discussed in Table 1.9 to Table 1.16 below.

Table 1.9: Key issue 8. Population: Study 301 and Study 303 excluded patients with mental health problems

Report Section	**4.2.3**
Description of issue and why the EAG has identified it as important	Studies 301 and 303 which inform the health economic model excluded patients with mental health problems. Because insomnia is frequently comorbid with other mental health problems the exclusion of patients with mental health problems may decrease the generalisability of the underlying evidence to the decision problem.
What alternative approach has the EAG suggested?	New evidence for patients with comorbid mental health problems receiving different treatments has to be collected.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	New evidence for patients with comorbid mental health problems receiving different treatments has to be collected.
EAG = Evidence Assessment Group	

Table 1.10: Key issue 9. Intervention and comparators: The company implemented only "no treatment" as a comparator

Report Section	**4.2.4**
Description of issue and why the EAG has identified it as important	A variety of pharmaceuticals and therapies are available for the treatment of insomnia. The company only included no-treatment as a comparator to daridorexant in the health economic model.
What alternative approach has the EAG suggested?	The EAG suggested in the clarification request that the company include relevant comparators such as sleep hygiene advice, CBT-I, non-benzodiazepine hypnotic medication, zolpidem, zopiclone, benzodiazepines and melatonin. The company did not comply with this request.
What is the expected effect on the cost effectiveness estimates?	Unclear, potentially small.
What additional evidence or analyses might help to resolve this key issue?	The health economic model has to include all relevant comparators.
CBT-I = insomnia-related cognitive behavioural therapy; EAG = Evidence Assessment Group	

Table 1.11: Key issue 10. Intervention and comparators: The 25 mg dosage was not included in the cost effectiveness model

Report Section	4.2.4
Description of issue and why the EAG has identified it as important	The company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation.
What alternative approach has the EAG suggested?	As evidence for the subgroup for which the 25 mg dosage would be used is missing, analyses are currently not possible.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	New evidence for the subgroup for which the 25 mg dosage of daridorexant would be used has to be collected. Meanwhile a scenario analysis using only data from patients who received the 25 mg population in studies 301 and 303 would be of interest.
EAG = Evidence Assessment Group	

Table 1.12: Key issue 11. Treatment effectiveness and extrapolation: Dropout adjustment

Report Section	4.2.6
Description of issue and why the EAG has identified it as important	As per the CS, the no-treatment arm was modelled to have no dropout, as patients receiving could not dropout from receiving no treatment. However, in the economic model provided by the company, the dropout rates observed in studies 301 and 303 for the daridorexant arm were applied to both daridorexant and no-treatment groups. This contradicts the statement made by the company.
What alternative approach has the EAG suggested?	Applying no dropout rates for the no-treatment arm, and the dropout rates from studies 301 and 303 to the daridorexant arm.
What is the expected effect on the cost effectiveness estimates?	Daridorexant is dominated by no-treatment.
What additional evidence or analyses might help to resolve this key issue?	N/A
CS = company submission; EAG = Evidence Assessment Group; N/A = not applicable	

Table 1.13: Key issue 12. Treatment effectiveness and extrapolation: Assuming no improvement in the no-treatment arm after the third month

Report Section	4.2.7
Description of issue and why the EAG has identified it as important	For the company base case placebo effect was only included for the first three months in the no-treatment arm, but not for the remaining 40 weeks. The EAG considers that the effect of selective attrition on the daridorexant group and the possibility of regression to the mean on the no-treatment group, were not sufficiently justified by the

Report Section	4.2.7
	company, and these effects could have biased the comparison in favour of the intervention.
What alternative approach has the EAG suggested?	The EAG suggested selecting the pessimistic scenario from the CS and applying the same placebo effect observed in both studies (301 and 303).
What is the expected effect on the cost effectiveness estimates?	The expected effect would be an increase in effectiveness on the comparator arm (no-treatment) and hence an increase on the ICER (from [redacted] to £[redacted])
What additional evidence or analyses might help to resolve this key issue?	N/A
CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; N/A = not applicable	

Table 1.14: Key issue 13. Adverse events exclusion from the cost effectiveness model

Report Section	4.2.7
Description of issue and why the EAG has identified it as important	The company excluded the AEs reported in studies 301 and 303 from their cost effectiveness model, assuming that these are minor AEs and would not be expected to have consequences on resource use or HRQoL.
What alternative approach has the EAG suggested?	An updated cost effectiveness model and scenario analyses incorporating all AEs from studies 301 and 303.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	An updated cost effectiveness model and scenario analyses incorporating all AEs from studies 301 and 303.
AEs = adverse events; EAG = Evidence Assessment Group; HRQoL = health-related quality of life	

Table 1.15: Key issue 14. Health-related quality of life: mapping of ISI© scores to EQ-5D utilities

Report Section	4.2.8
Description of issue and why the EAG has identified it as important	There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.
What alternative approach has the EAG suggested?	Scenario analysis incorporating a re-estimated mapping function in line with ISPOR Good Practices for mapping studies and including relevant covariates. Scenario analyses exploring ALDVMM and CLAD models. Detailed responses to all aspects/considerations mentioned in Tables 1, 2 and 3 of the ISPOR Good Practices for mapping studies.

Report Section	4.2.8
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Scenario analyses exploring ALDVMM and CLAD models and validation of the mapping function. Detailed responses to all (!) aspects/considerations mentioned in Tables 1, 2 and 3 of the ISPOR Good Practices for mapping studies.
ALDVMM = Adjusted Limited Dependent Variable Mixture Model; CLAD = Censored Least Absolute Deviations; EAG = Evidence Assessment Group; EQ-5D = European Quality of Life-5 Dimensions; ISI© = insomnia severity index; ISPOR = The Professional Society for Health Economics and Outcomes Research	

Table 1.16: Key issue 15. Resource use and costs: not all potentially relevant costs included in the economic model

Report Section	4.2.9
Description of issue and why the EAG has identified it as important	In addition to treatment acquisition costs, the CS only incorporated costs and resource use for GP visits, emergency room visits and inpatient care. The company justified the decision due to these being the only categories captured in the NHWS dataset and stating that the approach was conservative. Such a conclusion cannot be drawn in the absence of supporting evidence.
What alternative approach has the EAG suggested?	The EAG would prefer all costs relevant to the NHS and PSS perspective were included. NHWS data could be supplemented with alternative sources to inform costs that are currently not included.
What is the expected effect on the cost effectiveness estimates?	The EAG is unable to comment on the expected direction of impact on the ICER of including additional cost categories. However, doing so would provide a more accurate representation of the costs associated with the treatment and comparator in clinical practice.
What additional evidence or analyses might help to resolve this key issue?	Identification and inclusion of all additional cost categories, relevant to the NHS/PSS perspective, into the economic model.
CS = company submission; EAG = Evidence Assessment Group; GP = general practitioner; ICER = incremental cost effectiveness ratio; NHS = National Health Service; NHWS = National Health and Wellness Survey; PSS = Personal Social Services	

1.6 Summary of the EAG's view

The NICE scope and decision problem involved evaluation of daridorexant against established clinical practice. However, the company only provided evidence for daridorexant against placebo, without any attempt to compare daridorexant to established practice using indirect treatment comparisons. It is therefore difficult to clinically evaluate daridorexant in the appropriate context of the decision problem.

The evidence submitted lacked a key paper and was therefore incomplete. The two included studies suggest that daridorexant yields clinical benefits compared to placebo in the short term (3 months) but that in the longer term these benefits may become less certain. The EAG accepts that the uncertainty may be partly due to the lower statistical power of the longer-term study, but it cannot be assumed that this is the sole cause. Adverse events (AEs) appeared to be generally non-serious, and therefore less likely to have a significant negative impact on any benefits of daridorexant.

In terms of applicability, questions remain about the relevance of the study findings to the UK population. Although uncertain, there was a possible difference in the proportions of ethnicity groups between the target UK population and the two included studies. There is additional uncertainty about whether ethnicity is an important factor influencing outcomes: Study 301 did not sub-group for ethnicity, and whilst Study 303 did not find evidence that ethnicity was an effect modifier, analyses were only presented for two outcomes. Although there is no clear evidence that ethnicity is an effect modifier, there is insufficient evidence to support the company's claim that ethnicity is not an effect modifier. In addition, the study populations were largely naïve to the main alternative treatment insomnia-related cognitive behavioural therapy (CBT-I). This creates a serious divergence from the intended clinical population for daridorexant: people who have not responded to CBT-I.

The CS base case probabilistic and deterministic ICERs (with dropout adjustment) were [redacted] and [redacted] per QALY gained, respectively. The EAG base case probabilistic and deterministic ICERs, based on the EAG preferred assumptions highlighted in Section 6.1, were [redacted] and [redacted] per QALY gained. The most influential adjustment was related to the company's placebo correction for the no treatment arm. The ICER increased by [redacted] in the scenario assuming alternative dropout rates.

Remaining uncertainty about the effectiveness and relative effectiveness of daridorexant can be at least partly resolved by the company by conducting further analyses (as highlighted in Table 6.1) and providing further justification regarding the appropriateness of the mapping function. Moreover, the current assessment does not provide an appropriate estimation of the comparators listed in the scope.

Table 1.177: Probabilistic EAG base case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case (without dropout)*					
Daridorexant	[redacted]	0.724			
No-treatment	£637	0.691	[redacted]	0.034	[redacted]
CS base case (with dropout)*					
Daridorexant	[redacted]	0.543			
No-treatment	£478	0.518	[redacted]	0.024	[redacted]
EAG base case					
Daridorexant	[redacted]	0.720			
No-treatment	£622	0.703	[redacted]	0.017	[redacted]

* These results are slightly different from the ones stated in the CS, due to:
- The Excel model code
- The EAG has calculated the ICER from the total costs and QALYs from the PSA, not the incremental results from those.

CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG Comment
Population	Adults with insomnia disorder	Adults with insomnia disorder	N/A	There may be a mismatch between trial and UK patient populations. This may have implications for applicability. Although daridorexant is designed as a replacement treatment for those people that may be unsuitable for established care treatments such as CBT-I, most of those in the trials had never had the opportunity to receive or reject CBT-I. Again, this may have implications for applicability.
Intervention	Daridorexant	Daridorexant	N/A	In the CS the only dose that is considered is 50 mg, and analyses in the evidence involving 25 mg are not included. The company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation. This does not tally with the NICE scope that does not specify 50 mg. This specification is not justified. Furthermore, the duration of treatment and stopping rules are stated but not explained.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG Comment
Comparator(s)	Established clinical management (including sleep hygiene advice) without daridorexant	Established clinical management (including sleep hygiene advice) without daridorexant	N/A	In the CS the comparator is placebo, and not established clinical management. No indirect treatment comparison is used to attempt to rectify this issue. Therefore, there is a major difference between the decision problem comparator and the comparator in the evidence.
Outcomes	The outcome measures to be considered include: • Resolution of symptoms • Changes in sleep patterns and architecture • Sleep quality • Daytime alertness • Recurrence of insomnia • Adverse effects of treatment (including residual daytime sedation and memory impairment) • HRQoL	The outcomes addressed in this submission include: • Improvement of night-time symptoms of insomnia • Changes in sleep architecture and sleep efficiency • Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function • Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score • Rebound insomnia	Resolution of symptoms is not an appropriate term to describe the outcome in this submission. The outcome studied in the clinical trials of daridorexant is the quantitative and qualitative improvement of symptoms rather than resolution.	'Resolution of symptoms' is missing, and the justification is inadequate. The outcomes presented by the company do not necessarily fit into the NICE scope categories. No outcome data appear to be provided for some of the NICE scope outcome categories. Most importantly, there are a multiplicity of outcomes covering the same construct, which could increase the risk of type I errors.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG Comment
		• Adverse effects of treatment (next-day residual treatment effects and memory impairment) • HRQoL		
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and PSS perspective.	The cost effectiveness of daridorexant is presented as cost per QALY. Clinical and cost effectiveness of the reference case is estimated over a 12-month time horizon.	A short-term model estimating clinical and cost effectiveness over a 12-month time horizon is presented as the reference case for several reasons. Pharmacodynamics and clinical data of daridorexant demonstrate that the effect of treatment on sleep parameters occurs from the first day of treatment and that the effects on the sleep parameters are mostly lost on the first day of treatment discontinuation. In addition to presenting clinical and cost effectiveness over a 12-month time horizon, lifetime effects and potential QALY gains from better sleep (e.g., cardiac benefits, reduced fall risk, mortality) is discussed qualitatively in the submission. The potential quantitative impact of having a lifetime model, including impact of	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG Comment
			improved sleep duration on mortality and the impact of discontinuation, is presented as a scenario.	
Subgroups to be considered	The availability and cost of biosimilar and generic products should be considered. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	None	N/A	Sub-grouping has been carried out, but this appears to have been carried out arbitrarily, with some sub-grouping variables applied to some studies/outcomes but not to others. This makes it difficult to evaluate applicability. For example, for ethnicity (where a difference exists between UK population and the trials), ethnicity was not applied as a sub-grouping criterion to any of the analyses in Study 301 and to only two outcomes in Study 303. Therefore, ethnicity cannot be excluded as a potential covariate.
Special considerations including issues related to equity or equality	None specified.	None identified.	N/A – in line with the NICE final scope.	

Based on Table 1 of the CS[1]

CBT-I = insomnia-related cognitive behavioural therapy; CS = company submission; EAG = Evidence Assessment Group; HRQoL = health-related quality of life; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; N/A = not applicable; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom

2.1 Population

2.1.1 Definitions

The population defined in the scope is: *"Adults with insomnia disorder"*, which agrees with the stated population in the decision problem.

EAG comment:

- The definitions of this population are unclear in the company submission (CS).[1] In Table 2 of the CS the indication for daridorexant is "*adult patients with insomnia characterised by symptoms present for at least 3 months and considerable impact on daytime functioning"*.[1] Meanwhile, Section B.1.3 of the CS states that *"chronic insomnia, also known as insomnia disorder, is defined as symptoms occurring for ≥3 nights per week for ≥3 months together with daytime impairment"*.[1] These varying definitions, whilst not contradictory, suggested an inconsistent level of detail in defining the condition, leading to ambiguity in the decision problem definition.
- In the clarification letter, the company was asked to confirm the definition of the population.[2] In response, the company defined the population as follows: *"The population specified in the decision problem is adults with insomnia disorder. This is based on the definition provided by Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®), which defines insomnia disorder as 'dissatisfaction with sleep quantity or quality, associated with difficulty initiating or maintaining sleep, or early morning awakening. Furthermore, the sleep disturbance is associated with significant social or functional distress or impairment. Sleep difficulty occurs at least 3 nights per week and is present for at least 3 months and occurs despite adequate opportunity for sleep'. Additionally, the DSM-5® criteria of insomnia disorder is largely consistent with the patient population indicated in the Summary of Product Characteristics (SmPC) for daridorexant, and the same DSM-5® criteria has been used to enrol patients in the pivotal trials of daridorexant (studies 301 and 302)"*.[3] The EAG notes that this definition merges the two definitions reported in the CS, and can be taken as the more comprehensive and useful, definition.
- The company was also asked to define the typical symptoms of insomnia and to define daytime impairment.[2] In response, the company defined typical symptoms and daytime impairment as follows: *"The symptoms of chronic insomnia include problems of sleep initiation or maintenance despite adequate opportunities or circumstances of sleep which impacts daytime functioning. For diagnosis of insomnia disorder, current diagnostic classifications, viz. DSM-5® and International Classification of Sleep Disorders, Third Edition (ICSD-3) not only include symptoms of sleep difficulties, but also complaints of significant distress, or daytime impairment. Since insomnia disorder is a subjective condition, its diagnosis solely depends on patients' experience of sleep difficulties and daytime impairment. The common symptoms of distress due to daytime consequences include somnolence, fatigue, daytime sleepiness, cognitive deficit, mood disturbance, reduced motivation, proneness for accidents, and impaired work or relationship functioning. These symptoms may serve as primary indicators of daytime functioning impairment in clinical practice. Further, various patient-reported outcome instruments validated in clinical practice are available to assess patients' sleep habits and daytime functioning impairment. This includes: Daytime Insomnia Symptom Scale (DISS), the Daytime Consequences of Sleep Questionnaire (DCSQ), the Functional Outcomes of Sleep Questionnaire (FOSQ), the Pittsburgh Insomnia Rating Scale (PIRS), the Profile of Mood States (POMS), the Sleep Functional Impact Scale (SFIS), the Insomnia Severity Index (ISI), the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) and the Epworth Sleepiness Scale (ESS). As noted in our submission*

clinical guidelines do not recommend the use of any specific PRO [patient-reported outcome] *instrument to assess insomnia symptoms in clinical practice"*.[3] The EAG appreciates the clarity of this response and notes that this provides a far clearer picture of the disorder.

- The inclusion criteria for Study 301 contains more specific details that could be argued to make the population of the trial narrower than the population defined in the decision problem. For example, the trial inclusion criteria includes an ISI© score ≥15; sleep disturbance causing clinically significant distress or impairment in social, occupational, educational, academic, behavioural, or other important areas of functioning; self-reported insufficient sleep quantity (≥30 minutes to fall asleep, wake time during sleep ≥30 minutes, and subjective total sleep time (sTST) ≤6.5 hours during the night) for at least 3 nights per week during at least 3 months prior to the screening visit, and for at least 3 out of 7 nights on the Sleep Diary Questionnaire (SDQ) completed during the placebo run-in period prior to the run-in polysomnography (PSG) nights. Given the possibility that the population in the trials is narrower than the population in the decision problem, there are implications that the trial results might not necessarily be applicable to the clinical population. The company was therefore also asked to discuss applicability implications if the population in the decision problem turns out to be broader than that defined in Study 301.[2] The company responded by stating that *"the population in the decision problem (i.e., adults with insomnia disorder as per the DSM-5® criteria) is not expected to be broader than that of study 301. The use of ISI© score ≥15 as an inclusion criterion in study 301 is unlikely to impact the generalisability of the findings to the population in the decision problem since ISI© <15 represents subthreshold insomnia"*.[3] Despite the fuller definitions of chronic insomnia provided by the company, uncertainty persists because there remain some inclusion criteria for the trial (other than ISI© score) that are not covered precisely by the definitions of the conditions provided by the company (i.e. ≥30 minutes to fall asleep). Therefore, the EAG is still not fully convinced of the applicability of the trial results to the United Kingdom (UK) population.

2.1.2 Line of therapy of population

In the CS, it is stated that *"while digital or face-to-face CBT-I is recommended as the first-line treatment for insomnia disorder, it may not be suitable for or accessible to all patients…..up to 40% of patients refuse CBT-I, or cannot access it, when recommended by their general practitioners (GPs)"*.[1] Among those who receive either face-to-face or digital CBT-I, *"[redacted] fail to achieve the desired results, leading to an overall CBT-I success rate of only [redacted]"*. It is concluded that *"daridorexant may thus be suitable for this group of patients as an alternative first-line treatment"*.[1]

EAG comment:

- The data cited above are based on 'data in file', which does not appear to be peer-reviewed research material. In the response to the request for clarification, the company have been asked to provide the characteristics of the patients failing to respond to CBT-I, along with information to explain these values.[2] The company responded by stating that *"the CBTi refusal and failure rates were obtained from a recent survey conducted among 1,002 GPs in the UK. Respondents were asked up to 12 questions regarding insomnia; this included the number of insomnia patients seen in the last 3 months, standard insomnia treatment algorithms for patients with insomnia disorder, availability and funding of CBTi, its refusal and failure proportions and referral to secondary care. No patient characteristics were collected as part of the survey. Moreover, the NICE's assessment of Sleepio® highlighted the high dropout and failure rates with digital or face-to-face CBTi, which mentioned that the dropout rate was as high as 61.6%. This translates to a maximum success rate of 38.4%,*

which is close to the [redacted] *reported in the GP survey presented in the CS"*.[3] The EAG has not seen a copy of this report and so cannot evaluate it further.

- On reviewing the populations in studies 301 and 302 (a feeder trial of Study 303), it is apparent that the populations have had minimal exposure to CBT. For example, in Study 301 only 0.3% of participants were receiving cognitive behavioural therapy (CBT) at screening, 2.7% reported a previous failed CBT, and 87.9% of patients did not know CBT existed or were never offered CBT as a treatment option. The percentage of participants who had no access, interest or who refused CBT was 9.8% for all reasons combined. This has implications for applicability: if the target population for daridorexant are those who are non-responders to CBT-I, but the evidence has been yielded from all with chronic insomnia, then there may be differences in outcome between the trials and the real-world.
- In the request for clarification, the company has been asked to comment on the appropriateness of using a largely CBT-naïve population to justify the use of a pharmacological intervention as an alternative to CBT when it is apparent that most participants have never had the opportunity to receive or reject CBT.[2] The company stated that: "w*hile CBTi is the recommended first-line treatment for insomnia disorder, it is associated with certain limitations that bottlenecks its utilisation.*

    - *Poor access and availability of face-to-face CBTi has been a longstanding problem.*
    - *CBTi is resource intensive, and depending on the patient's need the number of sessions may vary from 6-8.*
    - *Adherence to CBTi is often poor as patients have to invest personal time and discipline to practise CBTi measures during and after the sessions.*
    - *Inconsistent results arise from lack of standardised accredited training for resources administering CBTi.*

    These limitations lead to high refusal and failure rates with CBTi, which may be reflective of the population in study 301. In such cases, clinicians resort to alternative pharmacotherapies (benzodiazepines, Z-drugs, and melatonin) for immediate relief of insomnia symptoms. As described in the CS, hypnotics can effectively treat night-time symptoms of insomnia disorder such as sleep onset and/or sleep maintenance, but psychological dependence often leads to its prescription longer than their recommended duration as no long-term alternates exist in clinical practice. NICE's recommendation for Sleepio® (a digital self-help CBTi for the treatment of insomnia disorder) may significantly improve the limitations of access and cost with CBTi, but as highlighted by NICE there is limited clinical evidence to show the effectiveness of Sleepio® compared with face-to-face CBTi".[3]

- The EAG is not satisfied with this response, because most participants in the trial never had the opportunity to receive or reject CBT. Therefore, they were not receiving daridorexant as a second line treatment. These were therefore not necessarily the same patients that would receive daridorexant in the real world.
- Finally, the company was asked if the population in the decision problem includes patients for whom CBT-I is not suitable or not accessible, as this is unclear.[2] The company response was as follows: "*In the decision problem, the population for daridorexant treatment includes patients for whom CBTi is inaccessible, unavailable or unsuitable i.e. as an alternative treatment. In addition, daridorexant may be used as second-line treatment, maintenance treatment, or for rapid symptom relief:*

- *For treatment-naïve patients who fail to respond to digital or face-to-face CBTi, daridorexant may be administered as a second-line treatment.*
- *For treatment experienced patients who have already completed standard of care including pharmacotherapy, daridorexant can be an alternative option.*
- *When longer-term management of insomnia symptoms (i.e., beyond 4 weeks) is required, daridorexant may be administered as maintenance treatment.*
- *When a patient is awaiting access to CBTi or a sleep specialist, daridorexant may be administered to provide rapid symptom relief"*.[3]

- The EAG is not satisfied with this response because, to reiterate previous arguments, the populations have had minimal exposure to CBT. For example, in Study 301 only 0.3% of participants were receiving CBT at screening, 2.7% reported a previous failed CBT, and 87.9% of patients did not know CBT existed or were never offered CBT as a treatment option. The percentage of participants who had no access, interest or who refused CBT was 9.8% for all reasons combined[3]. This does not sound like a population for whom "*CBTi is inaccessible, unavailable or unsuitable".*

2.1.3 Comorbidities

On page 25 of the CS, it is stated that *"multiple psychiatric and medical conditions are frequently associated with insomnia and may have a reciprocal relationship"*.[1] It is further stated that *"approximately 50% of patients with insomnia also have mood (e.g., major depressive disorder) or anxiety disorders (e.g., PTSD)"* [post-traumatic stress disorder]. The NICE final scope also states that *"insomnia is associated with comorbid conditions such as chronic obstructive pulmonary disease, heart failure, chronic pain, and psychiatric conditions (depression, anxiety, substance abuse, and post-traumatic stress disorder"*.[4] However, studies 301 and 302 (a feeder into Study 303) both exclude patients *"with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse…"* and Study 303 excluded those with *"ECG* [electrocardiogram] *findings"*, meaning those with cardiac issues may have been excluded.

EAG comment:

- These conflicts between comorbidities and exclusions lead to concerns about inappropriate exclusions from the trials, which may affect applicability.
- In the clarification letter, the company has been asked to comment on what extent these selection criteria might restrict the generalisability of the trial populations to the chronic insomniac population at large, and in England and Wales specifically. The company responded by stating that: "*The strict inclusion/exclusion criteria allowed the selection of a well characterised insomnia population, in need of pharmacological intervention, thus being representative of insomnia disorder population. The company acknowledges that many patients in clinical practice are likely to have comorbidities, including neuropsychiatric disorders resulting in the use of various concomitant CNS-active medications; however, the need to exclude subjects with some comorbid conditions was driven by the importance of limiting factors that could interfere with the optimal evaluation of the efficacy and safety of daridorexant. Since the underlying mechanisms of insomnia are thought to be the same in subjects with and without psychiatric disorders, including depression, the exclusion of these subjects does not affect the generalisability of the study results to insomnia disorder population at large, as well as to the population in England and Wales.*" [3] The EAG understand excluding people with conditions that would make it impossible to take part in a research project, and appreciate that those excluded may have fallen into this category.

2.1.4 Sleep hygiene advice

The NICE final scope recommends that "sleep hygiene advice" should be attempted before continuing along the treatment pathway.

EAG comment:

- No details of sleep hygiene advice is provided in the CS. [1] In the clarification letter, the company have been asked to provide details on the sleep hygiene measures that had been previously tried in the trial populations. The company responded by stating that: "*The median time since insomnia diagnosis of all subjects in study 301 was 7.1 years. Therefore, it can be assumed that most subjects have attempted sleep hygiene advice prior to study enrolment. Information regarding sleep hygiene advice was not collected for the trial population of study 301, as it would be prone to recall bias given that sleep hygiene advice is usually attempted shortly after diagnosing insomnia disorder before continuing along the treatment pathway*."[3] The EAG is satisfied with this response.

2.2 Intervention

The intervention defined in the final NICE scope is "*daridorexant*", without any stipulation of dose. The intervention in the decision problem is stated in the CS[1] as being the same. However, the CS[1] seems to focus on those participants treated with 50 mg; that is, the clinical and cost effectiveness analyses only include the 50 mg dose. Table 2 in the CS[1] also states that "*the treatment duration should be as short as possible. The appropriateness of continued treatment should be assessed within 3 months and periodically thereafter*".

EAG comment:

- Again, the restriction of the trial population relative to the decision problem has implications for applicability. In the clarification letter, the company was asked to justify the dose. The company stated that, "*The decision problem population excluded patients treated with 25 mg once daily dosage of daridorexant as this dosage is only indicated for patients with moderate hepatic impairment or where there is co-administration of moderate CYP3A4 inhibitors*."[3] The EAG considered this a reasonable justification but also thought that this presented an important problem for population applicability: the results from the trial are not applicable to people with moderate hepatic impairment or those on moderate cytochrome P450 3A4 (CYP3A4) inhibitors as they would be unable to use the 50 mg dose.

In relation to a question about the expected treatment duration and whether this could be longer than the Study 303 duration as well as the model time horizon of 12 months, the company stated that: "*The currently recommended drug classes for insomnia disorder are indicated for only a short duration (<4 weeks for hypnotics, ≤13 weeks for melatonin). However, in clinical practice these drug classes are commonly used beyond their recommended duration. A UK insomnia market landscape analysis showed that, on average patients were on prescription drugs for* [redacted] *days in 2021. Specifically, the average duration of therapy was* [redacted] *days for zopiclone,* [redacted] *days for melatonin and* [redacted] *days for amitriptyline. Given the chronicity of insomnia disorder, the treatment duration of daridorexant will likely be similar to or longer than these prescription drugs. Thus, the cost effectiveness model estimates ICER for the full population over the first 12 months and those remaining on treatment after 12 months (lifetime scenario)".*[3] The EAG considers that response, which focusses on other drugs, does not reduce the uncertainty around the optimal treatment duration for daridorexant.

- In relation to a question on stopping rules, the company responded as follows: "*With daridorexant, no formal stopping rules have been contrived. Per the SmPC the appropriateness of continued treatment should be assessed within 3 months of starting daridorexant and periodically thereafter. Primary care clinicians can monitor patient response and evaluate the need to continue treatment using established tools and approaches. Daridorexant's characteristic feature of quick onset and short half-life allows treatment benefit to occur rapidly while on the medication; however, treatment effect stops when treatment stops, as demonstrated by the placebo run-out phase in-between study 301 and 303. Patients who remain on treatment are likely to accrue the greatest treatment benefit*".[3] The EAG is satisfied with this response.

2.3 Comparators

The comparator defined in the scope is "*Established clinical management (including sleep hygiene advice) without daridorexant*".[4] The comparator in the decision problem is stated to be the same, but it is referred to as 'no treatment' in the cost effectiveness analysis.

EAG comment:

- However, this is questionable because in the trials evidence presented in the CS[1] the comparator is placebo, and not described as established clinical management. It may be noted that concomitant treatments in the trials were allowed alongside the randomised treatments. CBT-I was allowed provided it had been started 4 or more weeks prior to baseline and continued throughout the studies. Non-prohibited drugs that were part of the patients' normal care were also permitted. The level of CBT-I use in both arms in study 301 was very low (1 person in each group) but matched between arms. The usage of CBT-I in study 303, and the actual use of non-prohibited concomitant drugs in either study was not presented by the company. However, it can be assumed that because the study was double-blinded any concomitant treatments should have been comparable between groups; blinding would ensure there could be no way in which preferential provision could be administered. Therefore, any ECM used in the trials should have been comparable between groups and so it could be regarded as a comparison of daridorexant plus ECM vs. ECM, which does not equate to daridorexant versus ECM (the comparison apparently defined in the NICE scope). In fact, the company argue in the cost-effectiveness section of the CS that "*…the placebo arm of the trial serves as a proxy for no treatment*" (p. 108) .
- Therefore, there is a serious divergence between the scope and the evidence in terms of the comparators used. This will have an important impact on interpretation of evidence and makes it very difficult to form useful conclusions relevant to the decision problem.
- The first line clinical management for insomnia disorder is CBT-I, and, in line with the arguments above, it is unclear why this is not included as part of an ECM comparator in the decision problem.
- In the clarification letter, the company has been asked to explain why CBT-I was not included in the comparator. In response to this the company has stated that: "*In study 301, CBTi was not a feasible comparator considering the study's randomised double-blinded design. This design was necessary to minimise the impact of confounders and effect modifiers when assessing the efficacy and safety of daridorexant. However, CBTi was allowed as a previous or concomitant therapy. As a concomitant therapy, CBTi was allowed only if it was initiated at least one month prior to Visit 3, wherein the subject agreed to continue CBTi throughout the study. In clinical practice and in line with available guidelines, CBTi is recommended and should be offered as a first-line treatment for patients with insomnia disorder. However, in cases where digital or face-to-face CBTi is inaccessible, or where a patient is unable to follow CBTi steps, or refuses CBTi, daridorexant may be considered as an alternative pharmacological treatment. Pharmacological therapy should be*

started after CBTi has been offered and therefore CBTi was not considered as a comparator of daridorexant. This was discussed in detail during scoping, with feedback from clinical experts and patient groups, resulting in the removal of CBTi as a comparator from the final scope. This was reconfirmed in the Decision Problem Meeting".[3] The EAG response to this is that the decision problem defined the comparator as 'established clinical management'. Therefore, the comparator should have included CBT-I, as this is the established form of management for insomnia, unless the population is those who cannot or refuse to receive CBT-I. It is true that randomising participants to daridorexant and CBT-I would have made patient and health care provider blinding impossible, but this does not justify the failure to use the correct comparator. Comparing daridorexant to no treatment or daridorexant plus ECM to ECM (essentially excluding CBT-I), rather than ECM (including CBT-I) will give a much more optimistic effect size, and it could be argued that this gives a far more spurious result than performance bias resulting from lack of blinding in a comparison between daridorexant and CBT-I. Comparing to no treatment or ECM without CBT-I is therefore inappropriate unless part of an indirect treatment comparison with no treatment or ECM excluding CBT-I as common comparator, or the population precludes CBT-I.

- The company has also been asked that if CBT-I is not a comparator, then given that it is first line treatment should the population in the decision problem be modified to 2nd line, after CBT-I.[2] The company stated that: "*According to the positioning of daridorexant specified in B.1.3.6, CBTi is the first-line treatment for insomnia disorder, and in these patients, daridorexant will serve as a second-line option if patients fail to respond to digital or face-to-face CBTi. CBTi should always be recommended as first-line treatment for insomnia disorder. However, considering issues with access or inability of patients to follow CBTi steps or if patients refuse CBTi, daridorexant may be administered as an alternative pharmacological treatment*".[3] If the company are suggesting daridorexant as a first line treatment to replace CBT, then it should be compared against CBT. If the company are suggesting it as a second line treatment, then it should include a second line population where first line treatments have been tried and failed/refused. In either case, the company have not done this.
- In addition, the company was probed on whether CBT-I should be used as a concomitant therapy in the event of it not being a comparator, and to compare the rate of use of CBT-I between Study 301 and clinical practice in the National Health Service (NHS) of England and Wales, as well as to discuss the implications of any discrepancy. The company responded by stating that: "*In study 301, CBTi was allowed as a concomitant therapy. Only three randomised subjects (0.3%; 1 subject in each treatment group) were treated with CBTi at screening. Of the 927 subjects (99.7%) not using CBTi at screening, 25 subjects (2.7%; 11, 7, and 7 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported previous treatment failure with CBTi, 10 subjects (1.1%; 1, 5, and 4 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported no access/no therapist where subject lives, and 59 subjects (6.4%; 15, 24, and 20 subjects [daridorexant 25 mg, 50 mg, and placebo, respectively]) reported no reimbursement for CBTi (16). This highlights that study 301 has insufficient data to support the use of daridorexant as a concomitant therapy to CBTi, since only 0.1% of subjects were on concomitant CBTi. This was reflected in the company's proposed positioning of daridorexant (Section B.1.3.6)"*.[3]
- Finally, the company has been asked to include CBT-I as a comparator in the clinical and cost effectiveness analyses. The company stated that: "*CBTi was not specified as a comparator in the final scope of the decision problem, and this was discussed and agreed at the Decision Problem Meeting. Therefore, CBTi is not included as a comparator in the CS as per the positioning of daridorexant stated in A8 (a) and A8 (b)*." [3] The EAG would state that given the lack of an indirect

treatment comparison analysis, the failure to compare daridorexant to an established clinical management option such as CBT-I indicates that the decision problem has not been addressed.

2.4 Outcomes

The NICE final scope lists the following outcome measures:[4]

- Resolution of symptoms
- Changes in sleep patterns and architecture
- Sleep quality
- Daytime alertness
- Recurrence of insomnia
- Adverse effects of treatment (including residual daytime sedation and memory impairment)
- Health-related quality of life (HRQoL).

Resolution of symptoms was not included in the CS trials and was replaced by "improvement in symptoms". This was justified by the company on the grounds that it was "not an appropriate term", but no further rationale was given.

EAG comment:

- The company's argument that 'resolution of symptoms' is an inappropriate term appears to rest on the assumption that chronic insomnia disorder is lifelong and unresolvable. However, this does not tally with published data[5, 6] which show that around 30-40% of people appear to achieve long-term resolution. Although the replacement outcome of "improvement in symptoms" is not completely inappropriate, as it would encompass resolution of symptoms, it is possible that it might provide a more favourable picture for daridorexant.
- The outcomes reported in the trials do not all fit clearly into the outcome categories of the NICE final scope list. For example, WASO, total sleep time (TST), and latency to sleep onset (LSO) do not immediately appear to belong to any single category. The CS[1] states that WASO and LSO are measures of symptoms, and that TST is a measure of sleep architecture, so they have been placed in these categories in the report, but it is not immediately obvious why this is so. Further justification was requested from the company in the clarification letter. The company provided a table (Table 2.2) in their response as follows:

Table 2.2: Definitions of outcomes

Outcomes used in the CS	Definitions of outcomes used in the CS	Outcomes listed in NICE final scope
Improvement of night-time symptoms of insomnia	WASO (sleep maintenance), LPS (sleep onset), subjective TST (sleep time)	Resolution of symptoms
Changes in sleep architecture and sleep efficiency	Time to fall asleep, number of awakenings during the night and duration of TST by sleep stage/quarter of the night, depth of sleep	Changes in sleep patterns and architecture
Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function	Quality of sleep, daytime alertness and ability to function as assessed by VAS	Sleep quality

Outcomes used in the CS	Definitions of outcomes used in the CS	Outcomes listed in NICE final scope
Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score	Daytime impact of insomnia on three dimensions: physical (sleepiness domain), cognitive (alert/cognition domain), and affective (mood domain)	Daytime alertness
N/A	Recurrence of insomnia was not assessed	Recurrence of insomnia
Adverse effects of treatment (next-day residual treatment effects and memory impairment)	Withdrawal symptoms, rebound insomnia, next-morning residual effect and daytime sleepiness	Adverse effects of treatment (including residual daytime sedation and memory impairment)
Indirectly by mapping ISI© to EQ-5D	No specific questionnaire for HRQoL. Combination of the patient-reported assessments of sleep quality (using a VAS), daytime functioning (the IDSIQ questionnaire), and insomnia severity (the ISI© questionnaire).	HRQoL

Based on Table 4 from clarification question response from company[3]
CS = company submission; EQ-5D = European Quality of Life-5 Dimensions; HRQoL = health-related quality of life; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index; LPS = latency to persistent sleep; N/A = not applicable; NICE = National Institute for Health and Care Excellence; TST = total sleep time; VAS = visual analogue scale; WASO = wake time after sleep onset

The EAG will use these definitions to order the results sections.

- No outcome data appear to be provided for the outcomes of recurrence of insomnia (NICE final scope), rebound insomnia (company list of outcomes), quality of sleep (company list of outcomes), depth of sleep (company list of outcomes), daytime alertness (company list of outcomes) and daily ability to function (company list of outcomes). In the clarification letter, the company has been asked to identify the exact locations of these data in the report or add these data if necessary. The company responded as follows: "*Results for all the listed outcomes in the NICE scope and the company list of outcomes are provided in the CS, Appendix F or Appendix M. The table below* [Table 2.3]*indicates the exact locations of the data in question. Please note:*
    - *Recurrence of insomnia (NICE final scope) was not directly assessed in the trial subjects who experienced a treatment effect but those who subsequently discontinued treatment.*
    - *In the clinical trials presented in the CS, health-related quality of life (HRQoL) was assessed via IDSIQ, and no other instruments were utilised. HRQoL for the CEM was derived indirectly using ISI® scores collected from the trials mapped to EQ-5D, as described."* [3]

Table 2.3.: Company list of outcomes and their corresponding location in the CS/Appendix M/Appendix F

Company list of outcomes	Page number and Table number of results (as reported in CS/Appendix M/Appendix F)
Rebound insomnia	Study 301: Appendix F, Section F.1.1.4, Table 6 Study 303: Appendix F, Section F.1.2.4, Table 12
Quality of sleep	Study 301: Appendix M, Section M.1.3, Table 1

	Study 303: Appendix M, Section M.1.4, Figure 3
Depth of sleep	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Daytime alertness	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Daily ability to function	Study 301: Appendix M, Section M.1.3, Table 1 Study 303: Appendix M, Section M.1.4, Figure 3
Based on Table 6 from clarification question response from company[3] CS = company submission	

- Page 20 of the CS lists ISI© as a tool to measure the global assessment of insomnia severity. However, page 111 of the CS states that that same tool was used to quantify health-related quality of life (HRQoL) in both studies 301 and 303, which is how it has been classified in this report. In the clarification letter the company was asked to provide HRQoL results using validated tools such as EQ-5D. The company responded as follows: "*In studies 301 and 303, HRQoL was not assessed directly with HRQoL instruments such as EQ-5D; instead the ISI© was used to assess and monitor insomnia severity at baseline and at various timepoints after administration of study treatment. The Cerner Enviza NHWS was utilised to develop a mapping algorithm (Section B.2.9, CS). As EQ-5D was not included in the clinical study, utility was captured indirectly through mapping from ISI© using the mapping algorithm*".[3]The EAG is interested in why validated tools such as EQ-5D were not used as a direct measure. If there was a good reason for not using them this should have been fully justified.
- A major flaw in the presentation of the CS[1] is the use of several outcomes covering a single scope outcome. This will increase the risk of type I errors and is liable to present a more favourable picture for daridorexant. In the clarification letter, the company has been asked to provide a prioritisation of the outcomes within each category of NICE final scope outcome, with a clear rationale. The company response is as follows: "*The ISI© should be prioritised among all the outcomes presented in the CS as it is the key effectiveness parameter of the economic model. Given the complexity of assessing treatment outcomes in insomnia disorder, it is challenging to prioritise all other outcomes within each category of the NICE final scope since all outcomes within a category should be considered in totality and therefore carry equal importance when evaluating the clinical benefit of daridorexant. This is supported by a number needed to treat (NNT) analysis of the key endpoints of study 301 (i.e., WASO, LPS, sTST, IDSIQ and ISI©). The results of the NNT analysis show that all key endpoints have comparable NNTs at month 3, as indicated by the overlapping confidence intervals* (Table 2.4)."[3]

Table 2.4: Number needed to treat values for the responder analysis based on sTST, LPS, WASO, ISI© and IDSIQ at Month 1 and Month 3 for daridorexant compared with placebo

Variable	Threshold response definition	1 month	3 months
		Daridorexant 50 mg NNT, mean (95% CI)	Daridorexant 50 mg NNT, mean (95% CI)
sTST	Change from baseline of ≥55 min	[redacted]	[redacted]
LPS	LPS <20 min	[redacted]	[redacted]
WASO	WASO <30 min	[redacted]	[redacted]

Variable	Threshold response definition	1 month	3 months
		Daridorexant 50 mg NNT, mean (95% CI)	Daridorexant 50 mg NNT, mean (95% CI)
ISI©	Change from baseline of ≤-7 points	[redacted]	[redacted]
	Total score ≤7 points	[redacted]	[redacted]
IDSIQ	Change from baseline in sleepiness domain score of ≤-8 points	[redacted]	[redacted]
	Change from baseline in sleepiness domain score of ≤-4 points	[redacted]	[redacted]
	Change from baseline in alert/cognition domain score of ≤-12 points	[redacted]	[redacted]
	Change from baseline in alert/cognition domain score of ≤-9 points	[redacted]	[redacted]
	Change from baseline in mood domain score of ≤-7 points	[redacted]	[redacted]
	Change from baseline in mood domain score of ≤-4 points	[redacted]	[redacted]
	Change from baseline in total score of ≤-25 points	[redacted]	[redacted]
	Change from baseline in total score of ≤-17 points	[redacted]	[redacted]

Based on Table 5 from clarification question response from company.[3]
CI = confidence intervals; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index; LPS = latency to persistent sleep; mg = milligrams; min = minutes; NNT = number needed to treat; sTST = subjective total sleep time; WASO = wake time after sleep onset

The EAG opinion on this is that the problem of multiple outcomes remains, regardless of the apparent equality of the many outcomes. The company has not prioritised the outcomes per construct and so the risk of type I errors persists.

2.5 Other relevant factors

According to the company's 'data on file':

- Approximately 3.3 million adults in England suffer from insomnia disorder, with a substantial impact on patients' QoL and productivity.
- Both face-to-face and digital CBT-I (e.g., Sleepio®) have high refusal and failure rates. Among patients who are eligible for CBT-I, only [redacted] achieve the desired results.
- None of the other commonly prescribed insomnia treatments in the UK fulfil the criteria of an ideal treatment.

- There is therefore a need for an evidence-based treatment for insomnia disorder that is safe and effective for longer-term use. This will have an immediate impact on patients' QoL and productivity.

Daridorexant is the first dual orexin receptor antagonist (DORA) to be approved in the UK and Europe for the treatment of insomnia disorder. It is an evidence-based treatment with established efficacy and safety for up to one year.

Anticipated marketing authorisation is unclear. In the clarification letter response, the company stated that: "*Currently, marketing authorisation approval of daridorexant is still pending for MHRA. In March 2021, marketing authorisation application for daridorexant was submitted to the EMA. A positive CHMP opinion was issued in February 2022, and marketing authorisation was approved on 29th April 2022 by EMA for "the treatment of adult patients with insomnia characterised by symptoms present for at least 3 months and considerable impact on daytime functioning (32)." The marketing authorisation by MHRA is anticipated to be consistent with that of EMA*.".[3]

Daridorexant is Food and Drug Administration (FDA)-approved (January 2022).[7]

This appraisal does not fulfil the end-of-life criteria as specified by NICE.

There appear to be no equality considerations, other than the fact that "*a broad recommendation for daridorexant to treat insomnia disorder in primary care will provide GPs with a safe and effective option for patients who refuse or fail CBT-I*" (CS, Section B.1.4).

EAG comment:

- The data on file provide justification for a new treatment, but the data have not been made available to the EAG.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company conducted a systematic literature review (SLR) of the published literature to identify evidence on the clinical efficacy and safety of daridorexant and relevant comparators in patients who were suffering from (chronic) insomnia disorder. This Section of the EAG report describes and critiques the methods of the review including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

3.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the CS.[1] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[8, 9] The CS[1] was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[10] The EAG has presented only the major limitations of each search strategy in the report.

Appendix D of the CS details the SLR undertaken to provide a comprehensive assessment of the current evidence from randomised controlled trials on the efficacy and safety of pharmacological treatments for insomnia disorder in adults.[11]

A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE and MEDLINE In-Process	Ovid	All*	1/3/22
Embase	Ovid	All*	1/3/22
CENTRAL	Ovid	All*	1/3/22
PsycINFO	Ovid	All*	1/3/22
Conferences			
British Sleep Society	Internet	Two most recent meetings	Not stated
European Sleep Research Society			
ISPOR			
ISPOR Europe			
Trials registries			
ClinicalTrials.gov	Internet	All years	Not stated
WHO ICTRP	Internet	All years	Not stated

* The CS and response to clarification state that no date limit was applied, however it is not clear which database segment was used as the database start and end dates were not supplied[1, 3]

CENTRAL = Cochrane Central Register of Controlled Trials; ICTRP = International Clinical Trials Registry Platform; ISPOR = International Society for Pharmacoeconomic and Outcomes Research; SFRMS = Société Française de Recherche et Médecine du Sommeil; WHO ICTRP = World Health Organization

EAG comment:

- Searches were undertaken in March 2022 to identify randomised controlled trials (RCTs) on the efficacy and safety of pharmacological treatments for insomnia disorder in adults. The CS, Appendix D and the company's response to clarification provided sufficient details for the EAG to appraise the literature searches.[1, 3, 11]
- A good range of databases and trials registers were searched. Reference checking was conducted on bibliographies of systematic reviews and/or meta-analyses of RCTs evaluating pharmacological treatments for insomnia disorder, identified through the electronic literature database searches, and published since January 2017.
- Database searches were not limited by publication date or by language.
- Conference proceedings searches were conducted for the two most recent meetings available for four named conferences. However, conference proceedings were excluded from the Embase search, which can be a useful source of additional conference papers. In response to clarification, the Company stated that: 'Conference proceedings were excluded from the Embase clinical effectiveness searches due to a high volume of yield resulting from any conference proceedings reporting on 'insomnia', introducing a high number of irrelevant publications to screen. Hence, a targeted approach was followed by specifically hand searching conferences of interest in the past two years. It is standard practice to search for conference proceedings of preceding two years, as any study results published before would be reported in a peer review publication, which can be captured through database search.'

  However, the exclusion of conference proceedings only removed around 800 references from the Embase results, so would not have greatly increased the screening burden. Amongst these results were references from the World Sleep Congress and the Annual Meeting of the Associated Professional Sleep Societies and more generic neurology conferences, which may have provided additional useful references. The EAG notes also that it is not necessarily the case that all conference proceedings will be published in peer reviewed journals.
- Searches were well structured, transparent and reproducible, and a good range of subject indexing terms (MeSH/EMTREE) and free text was used.
- Database search strategies contained a population facet for insomnia and sleep disorders and the relevant measurement tools. This facet was then combined with terms for daridorexant and other drug therapies. Unlike the cost effectiveness searches, the strategies did not include any search terms for cognitive behaviour therapy, so would not have identified any studies only on CBT for insomnia disorders.
- Study design filters to identify RCTs were applied to the searches of Embase, MEDLINE, MEDLINE In-Process, CENTRAL and PsycINFO. The study design filters were not referenced, so it was unclear whether the filters used were published objectively derived filters. The filters contained a combination of subject heading terms and free text terms and the EAG deemed them to be adequate, although additional terms could have been added to the filters to improve recall, such as 'randomized controlled trial.pt.' in the MEDLINE strategy, and the free text terms 'placebo' in the PsycINFO strategy and 'RCT' in all strategies. The EAG also notes the use of the RCT filter in the CENTRAL search. As CENTRAL is a trials database the EAG believes it was not necessary to include this filter in the strategy, and this may have resulted in unnecessarily restricting the results retrieved.
- Separate searches for safety outcomes were not conducted. Guidance by the Centre for Reviews and Dissemination (CRD)[12] and Golder et al.[13] recommend that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed. The EAG considers it possible that relevant

evidence from studies other than Study 301 and its safety and tolerability extension Study 303 may not have been identified as a consequence of the RCT study design filter used in the database searches.

3.1.2 Inclusion criteria

The eligibility criteria used in the systematic literature review (SLR) were included in Appendix D[11] of the CS and are presented in Table 3.2. It was apparent on reviewing the eligibility criteria that there were some concerns identified by the EAG, principally concerned with the populations and comparator definitions.

The population, intervention, comparator and outcome (PICO) used in the eligibility criteria listed 'placebo' and 'other active agent' as the viable comparators, with an exclusion of 'non-pharmacological interventions.' This criterion would therefore exclude the inclusion of any trial where there was a comparison against 'sleep hygiene' methods, or CBT, both of which are current and appropriate primary treatments in England and Wales, and which the NICE final scope recommends. Additionally, the NICE final scope[4] lists the viable comparators as Established Clinical Management (ECM; including sleep hygiene advice) without daridorexant, indicating a discordancy with the PICO provided by the company in their identification of evidence.

In the request for clarification, the EAG asked the company to provide justification for their choice of comparator.[2] In their response, the company stated that *"CBTi was not a feasible comparator considering the study's randomised double-blinded design. This design was necessary to minimise the impact of confounders and effect modifiers when assessing the efficacy and safety of daridorexant. However, CBTi was allowed as a previous or concomitant therapy. As a concomitant therapy, CBTi was allowed only if it was initiated at least one month prior to Visit 3, wherein the subject agreed to continue CBTi throughout the study".*[3]

The EAG does not consider that this has explained or justified the company decision to exclude 'nonpharmacological interventions'. While the EAG understands the comments regarding the impossibility of blinding, this does not justify the use of a different comparator and in fact this has explicitly removed what would be considered the established treatment (1) recommended in England and Wales, (2) is the appropriate first line treatment and (3) includes the appropriate therapy (CBT) that the company claims can be replaced by the use of daridorexant. As stated in Section 2.3 of this report, using placebo as the comparator will likely produce a more optimistic effect size.

Of note, in the company response, was the following '*However, in cases where digital or face-to-face CBTi is inaccessible, or where a patient is unable to follow CBTi steps, or refuses CBTi, daridorexant may be considered as an alternative pharmacological treatment.*'[3] This states that if CBT-I is inaccessible, unable to be followed, or is refused, then daridorexant can be considered. However, in Study 301 the overwhelming majority of patients (87.9%) had not had this opportunity. The company go on to say, '*Pharmacological therapy should be started after CBTi has been offered and therefore CBTi was not considered as a comparator of daridorexant'.*[3] Again, the EAG reiterate that most participants in Study 301 had never heard of or been offered CBT. This submission promotes the data from this trial as a justification to offer daridorexant as an alternative to CBT despite the patients not having had the offer of CBT, and despite the lack of any comparison against CBT.

Additionally, the company in their response to clarification stated that '*daridorexant is a pharmacological treatment positioned in second line after interventions such as sleep hygiene and CBTi. Hence, the comparators of interest for this SLR were placebo or active agent.*' [3]If daridorexant

was indeed positioned as a second line treatment, then the population of patients included should have been a second line population, however this was not the case.

The EAG were curious about the included population definition. While the PICO included adults who were suffering from chronic insomnia disorder, which was in line with the NICE final scope of adults with insomniac disorder and so was technically appropriate, the EAG did consider whether it would have been more relevant to include patients who had either been exposed to first line therapies of sleep hygiene and CBT but had experienced failures (i.e., second line) or had refused it. The EAG considered this an important point, given that (a) the company emphasised the unpublished data suggesting that [redacted] of patients are unwilling or unable to receive CBT, and of those who do [redacted] experience treatment failures, and (b) the company promoted the potential of daridorexant as an alternative first line therapy to those who are unable or unwilling to receive CBT[1]. The EAG therefore considers it relevant and justifiable that the population should include those where CBT as a first line therapy has failed (if daridorexant is a second line treatment) or where CBT has been refused/not been accessible, if daridorexant is being proposed as an alternative first line therapy (and should therefore be reflected with CBT as a comparator, as discussed above).

The EAG reviewed the clinical study reports (CSRs) of studies 301 and 302 and were surprised to see that the vast majority (87.9%) of the trial populations had not been offered or were not aware of CBT. In its request for clarification, the EAG asked the company to comment on the appropriateness of using a largely CBT naïve population to justify the use of a pharmacological intervention as an alternative to CBT when it is apparent that most participants have never had the opportunity to receive or reject CBT. In their response the company reiterated their position that CBT-I has various limitations and that '*These limitations lead to high refusal and failure rates with CBTi, which may be reflective of the population in study 301'*.[3] This is not a satisfactory response and does not address the question. Firstly, it is speculative to claim this 'may be' reflective of the population. Secondly, as the CSR for Study 301 states clearly and unambiguously, that 87.9% of patients did not know CBT existed or were never offered CBT as a treatment option, these limitations and failure rates are in no way 'reflective' of the trial population. The response makes further statements that when CBT-I is not available, various other pharmacological and non-pharmacological treatments may be prescribed instead. It is stated that '*Sleepio® (a digital self-help CBTi for the treatment of insomnia disorder) may significantly improve the limitations of access and cost with CBTi, but as highlighted by NICE there is limited clinical evidence to show the effectiveness of Sleepio® compared with face-to-face CBTi*'.[3] Again, this is not relevant to the question asked of the company, which seeks justification for use of a CBT naive population in a submission for daridorexant, which is suggested to be a pharmacological replacement therapy when CBT cannot be accessed or is refused, and the data from trial 301 shows clearly that most of the trial population did not have this opportunity.

The eligibility criteria used in the search strategy for RCTs and non-RCTs is presented in Table 3.2.

Table 3.2: Eligibility criteria used in search strategy

Criteria	Inclusion Criteria	Exclusion Criteria
Population	Adults ≥18 years old with a diagnosis of chronic insomnia disorder according to any standardised diagnostic criteria (e.g., DSM, ICSD, ICD).	Paediatric (<18 years old) patients. Patients without chronic insomnia disorder or patients with chronic insomnia disorder according to unspecified diagnostic criteria. Patients with short-term (acute) insomnia.

Criteria	Inclusion Criteria	Exclusion Criteria
Intervention	Individual pharmacological interventions: Benzodiazepines: brotizolam, clonazepam, diazepam, estazolam flunitrazepam, flurazepam, haloxazolam, loprazolam, lorazepam, LMZ, midazolam, nimetazepam, nitrazepam, quazepam, rilmazafone, TMZ, and TZ. Benzodiazepine-like agents (Z-drugs): ESZ, ZAL, ZPD, and ZPC. Antidepressants: amitriptyline, DOX, mirtazapine, and TRA. Melatoninergic drugs: MEL and RAM. Orexin receptor antagonists: SUV, DAR, almorexant, filorexant, and LEM. Other: triclofos sodium.	Non-pharmacological interventions. Barbiturates, chloral hydrate, ethchlorvynol and quetiapine. Herbal products and medical devices. Combination therapy (e.g., CBT-I + pharmacological treatment or augmentation studies [drug A plus drug B versus drug A]).
Comparison	Placebo. Another active agent.	Comparison of different doses/preparations of the same active agent. Non-pharmacological interventions.
Outcomes	Efficacy§ ISI© Sleep quality (PSQI, LSEQ or other relevant scales) Sleep quantity parameters: Sleep maintenance (WASO) Latency to persistent sleep TST SF-36 SSS ESS IDSIQ´s sleepiness domain score Safety AEs SAEs Discontinuations Withdrawal, rebound, tolerance and addiction	Any other outcome not listed in the inclusion criteria.
Study Design	RCTs (phase >I). SLRs and meta-analyses of RCTs (for citation-chasing only).	Any other study design, including: Case reports Case series Animal studies/models Pharmacodynamic/pharmacokinetic studies Observational studies Phase I trials (e.g., dose-finding, dose-escalation studies) Single-arm trials Non-randomised trials Cross-over RCTs not reporting data before cross-over.

Based on table 6 of appendix D, CS[1]

§ Both objective and subjective measures are of interest

AEs = adverse events; CBT-I = cognitive behavioural therapy for insomnia; CS = company submission; DAR = daridorexant; DOX = doxepin; DSM = Diagnostic and Statistical Manual of Mental Disorders; ESS =

Criteria	Inclusion Criteria	Exclusion Criteria
Epworth Sleepiness Scale; ESZ = eszopiclone; ICD = International Classification of Diseases; ICSD = International Classification of Sleep Disorders; IDSIQ = Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index; LEM = lemborexant; LMZ = lormetazepam; LSEQ = Leeds Sleep Evaluation Questionnaire; MEL = melatonin; PSQI = Pittsburgh Sleep Quality Index; RAM = ramelteon; RCT = randomised controlled trial; SAE = serious adverse event; SF-36 = 36-Item Short Form Survey; SLR = systematic literature review; SSS = Stanford Sleepiness Scale; SUV = suvorexant; TMZ = temazepam; TRA = trazodone; TZ = triazolam; WASO = wake time after sleep onset; ZAL = zaleplon; ZPC = zopiclone; ZPD = zolpidem		

3.1.3 Critique of data extraction

Appendix D of the CS provides clarity on the process of data extraction[11]. The company state that eligibility screening was conducted in two stages. During level 1 screening, titles and abstracts were reviewed independently by two researchers. Disagreements between the reviewers were resolved by a third reviewer. During level 2 screening, those articles deemed eligible during level 1 screening were reviewed independently by two researchers as full texts. Disagreements between the reviewers were resolved by a third reviewer, as needed. This represents the optimal process of screening and reduces the opportunity for error and bias.

While most of the process was generally reasonably described. The EAG noted that some further clarity was required to fully explain and describe the methods.

The company described their processes of data extraction essentially as consisting of data extraction by the first reviewer, validation by the second reviewer, and then disagreements resolved by a third reviewer. Some further detail could have been provided in the submission to emphasise the attempts to reduce error/bias. It would have been helpful for the company to explain how the second reviewer validated data and to explain if it was conducted independently of the first reviewer, and how disagreement resolution by the third reviewer took place. Clarity around whether the third reviewer independently extracted the data or simply reviewed the basis of disagreement would be helpful. It would also be helpful to know if the third reviewer had conversations with one/both/none of the two reviewers before making a decision. Clear and descriptive explanations of these points help to provide reassurances that processes were mitigated as much as possible to try and reduce any likelihood of error or bias. This in turn, provides more credibility to the data and its ultimate conclusions. The optimal method would have seen two reviewers extracting data independently of each other with any disagreements resolved by a third. This method reduces the likelihood of error or bias.

The company also describe the role of prioritisation by Evidera. In appendix D[11] it is stated that '*Evidera conducted a high-level comparability assessment. These trial characteristics did not lead to exclusion from the SLR, but instead identified trials that were less likely to connect to a network for later analysis. Trials that did not connect to a likely network for later analysis were de-prioritised: those trials that did not include a treatment arm using a licensed dose of a treatment of interest, those that did not report comparable data, or those that did not report data in populations most of interest. Upon determination of the trials that were most likely to be suitable for NMA, a subset of trials underwent full extraction'*

The EAG did not fully understand how this was concordant with the process of data extraction. This text seemed to suggest that only trials that were deemed '*most likely'* to be NMA suitable had underwent full data extraction. Two questions naturally arose from the EAG at this point, the first was by what method an article was deemed to be likely for network meta-analysis (NMA) suitability and what were the processes of assessment and agreement by reviewers. The second was whether this suggested that some articles that otherwise met the full eligibility criteria, did not undergo full data extraction if they did not appear to be appropriate for the '*subset*'. The EAG asked the company to verify that all trials,

which met the PICO criteria and were included at full screening stage, did undergo data extraction and the EAG sought clarification on the manner of the '*high-level comparability assessment that was conducted by Evidera'*.

In their response to clarification, the company confirmed *'that no trials were excluded from the results of the Evidera comparability assessment'* and that *'A top-level extraction was performed for all included studies. This top-level extraction – 2 step extraction – provided sufficient information and data to determine comparability assessment - information on trial, patient, treatment characteristics, and outcome availability (i.e., tagging for the presence of relevant outcomes) were recorded. Studies (as listed in Table 7, Appendix D) were then de-prioritised if they did not include a treatment arm using a licensed dose of a treatment of interest, if they did not report comparable data, or they did not report data in populations of most interest.'*[3]

To the EAG, this indicates that those studies which were initially included for full screening (those that had been deemed eligible for inclusion at level 1 title/abstract screening) underwent full text screening, and then those included articles (those deemed eligible for inclusion at level 2 full text screening) were assessed for comparability by Evidera and ranked according to priority. The company in their response referred to a '*2 step extraction'* which the EAG presumes to mean the recording of '*information on trial, patient, treatment characteristics, and outcome availability (i.e., tagging for the presence of relevant outcomes)'* as the first step used to prioritise studies, with full data extraction then being the second step. Appendix D in the submission states '*Upon determination of the trials that were most likely to be suitable for NMA, a subset of trials underwent full extraction',* while, as mentioned above, the company in their response to clarification state *'that no trials were excluded from the results of the Evidera comparability assessment'*. However, this information suggests that the comparability assessment (step 1) determined which studies ('*subset of trials')* were then subject to full data extraction.

With regard to the process of data extraction itself, and our observations detailed above, the company in their response to clarification stated that '*Once the extractions were validated, these were sent back to the researcher who had performed the original extractions to make required changes. Any disagreements between the extractor and validator were brought forward and were resolved by a third, more senior investigator who reviewed the disagreement and provided a final decision[3]'*.

While the EAG appreciates the further information, the process of 'validation' itself is where the EAG would have liked to have seen some further description, with particular to how the data was checked. Typically, validation is quite a broad definition but usually means that a second independent data extraction has not been conducted, and that the second reviewer has checked what has been completed by the first reviewer. This is obviously more prone to error and bias than duplicate data extraction by two independent reviewers. The inclusion of the third reviewer does provide some mitigation, although one issue with this process is that there is an increased likelihood that only data that is extracted onto the extraction sheet is validated with disputes resolved by the third reviewer, while other relevant data that may have been missed by the first reviewer and not extracted in the first instance, remains as such. The EAG is also unclear as to whether the process of data extraction described above applied to both steps of the '2 step' extraction that was described and given that the comparability assessment (step 1) conducted by Evidera determined the subset of studies that were to be full screened (step 2), fully described processes are helpful. The EAG emphasises that the reporting of SLR processes should be clear, unambiguous, and described with appropriate detail to install confidence that likelihood of error and bias has been reduced as far as possible.

3.1.4 Quality assessment

The company state in Appendix D of the CS that all RCTs were assessed using the Cochrane Risk of Bias Assessment Tool 2.0[11]. Table 19 in Appendix D of the CS lists the risk of bias (RoB) results of all trials that the company claim met their eligibility criteria. However, despite Study 303 being included within the CS as an extension trial of Study 301, it is not listed here, and no RoB assessment appears to have been included. Study 301 was assessed by the company and was deemed to be of a low RoB overall. The CS[1] states that '*One researcher extracted data from the included papers into the DET, which was then validated by a second, senior investigator*'. This is lacking in the appropriate level of clarity and reporting, and potentially highlights a process that was at elevated RoB.

3.1.5 Evidence synthesis

In Appendix D[11] of the CS, it is stated that '*A systematic literature review (SLR) was undertaken to provide a comprehensive assessment of the current evidence from randomised controlled trials (RCT) on the efficacy and safety of pharmacological treatments for insomnia disorder in adults, with the potential to conduct an NMA'*. The company also describe the 'prioritisation' of studies by virtue of a '*high level comparability*' assessment designed to identify '*trials that were less likely to connect to a network for later analysis'*. Although the intentions were to conduct a network meta-analysis if appropriate, the company only identified one trial (Study 301) in the SLR which they deemed to be appropriate for inclusion in the submission. Study 303 was not identified in the SLR and is an extension of Study 301. Therefore no (network) meta-analysis was conducted.

EAG comment:

- While screening appears to have been conducted in duplicate**,** The Cochrane Handbook for Systematic Reviews recommends that "*as a minimum, information that involves subjective interpretation and information that is critical to the interpretation of results (e.g. outcome data) should be extracted independently by at least two people"*[14] Due to one reviewer completing data extraction and one person validating, there is an increased risk of bias and/or errors.[14]
- Although the chosen method of data extraction (one reviewer completing data extraction and one person validating) is a widely used method. Further detail and description could have been provided to provide reassurance that all necessary steps were taken at all stages to reduce the likelihood or error/bias, but also to simply provide the optimal clarity of reporting that is essential to a well conducted and reported SLR.
- The two-step data extraction conducted appears to be in essence a further step where all articles which had been included after level 1 and 2 screening was completed, were then prioritised by Evidera to identify only some articles that should undergo full data extraction. The EAG would like to have seen further detail about this process, particularly regarding how the likelihood of error and bias was mitigated. The EAG notes that our comments here do not necessarily suggest error or bias, but rather the EAG feels the processes could have been described and explained with more detail.
- The EAG notes lack of clarity and specific focus with respect to daridorexant and the appropriate 'population' and 'comparator' elements of the PICO. If daridorexant is a proposed second line treatment, the EAG thinks the population should be second line and should include those who have had access to and failed first line treatments, however this was not the case. If daridorexant is being proposed as an alternative first line therapy to replace the standard treatment of CBT-I, then daridorexant should be compared against CBT, but this was not the case.
- It was not fully clear how the process of quality assessment was conducted. No additional information could be identified in the CS main document or in Appendix D, to describe the details of the full method of quality assessment. While the appropriate RoB tool was described it was not

clear how many reviewers were involved (was there a third reviewer to resolve disagreements?), if it was conducted independently (how was the extracted data validated? with or without communication with the first reviewer?) or how disagreements were resolved (the decision of the senior investigator or consulting an independent third?). Again, this lack of reporting detail means that the EAG must consider that there is an increased likelihood of error and bias, and the EAG emphasises the point that a well conducted SLR must also be well reported with sufficient detail and clarity.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

In the abstract/title screening phase of the CS SLR, 4,089 records were excluded and 812 were retained for full text screening. Although 102 publications reporting on 63 RCTs were included in the company's SLR, this covered many pharmacological approaches to managing insomnia. The only included trials relevant to the decision problem were two RCTs evaluating daridorexant versus placebo[15, 16] and one RCT comparing daridorexant versus ZPD versus placebo.[17] Details of all three relevant RCTs included in the SLR are given in Table 3.3.

Only one of these three trials, CSR 301,[15] which covered daridorexant versus placebo, was eventually included in the clinical effectiveness evidence in the CS. CSR 302[16] was not included in the clinical effectiveness evidence in the CS[1] as the doses were outside the standard 50 mg. Dauvilliers et al. 2020[17] was also not included, although the reasons for this were unclear.

The clinical effectiveness evidence in the CS[1] additionally correctly included trial CSR 303,[18] although this was *not* included in the SLR. Reasons for its absence from the SLR are also unclear.

Table 3.3: Trials included in the CS SLR

Trial	Treatment	Inclusion in CS	EAG comments
NCT03545191 (22-29) CSR 301[15]	DAR 25 mg, 50 mg PBO	Yes	Correctly included as a key study.
NCT03575104 (23-25, 27, 30, 31) CSR 302[16]	DAR 10 mg DAR 25 mg PBO	No	Excluded because of wrong dose.
NCT02839200 (32, 33) Dauvilliers Y 2020[17]	DAR 25 mg, 50 mg DAR 5 mg, 10 mg ZPD 10 mg PBO	No	Reasons for exclusion from CS unclear. The CS[1] claims it was a 'dose finding study' but this is not the case as there is a placebo comparison against 50 mg DAR. Appears to be a key study.
Based on Table 8, Appendix D of CS[11] CS = company submission; CSR = clinical study report; DAR = daridorexant; EAG = Evidence Assessment Group; mg = milligram; PBO = placebo; SLR = systematic literature review; ZPD = zolpidem			

EAG comment:

- The lack of Dauvilliers et al. 2020[17] in the CS[1] appears to be a major omission and the company was asked to explain this in the clarification letter. The company responded by stating that: "*As elaborated in Table 4 of CS, the clinical trial programme of daridorexant included two phase II studies, one of which was NCT02839200. The study was a 6-arm randomised trial, that included 4 dosages of daridorexant, zolpidem and placebo. Primarily, this trial assessed dose response*

relationship between 5, 10, 25 and 50 mg dose of daridorexant and thus, was not designed to evaluate efficacy and safety of daridorexant compared with placebo due to the small sample size utilised in the study. Therefore, this study was not found to be relevant for the appraisal". [3] The EAG is not satisfied with this response. The study compared placebo to 50 mg of daridorexant and therefore, according to the protocol, is eligible for inclusion. There were no exclusion criteria relating to study size. It therefore remains unclear why this study was omitted and the EAG would have liked to see the results of this study included.

- The Company was also asked to explain the omission of CSR 303[18] from the SLR. The company stated that, "*Study 303 did not meet the SLR requirements due to the study design issues. In this extension study, subjects who had completed the study treatment and run-out period for studies 301 and 302 were re-randomised to receive either placebo or 25mg daridorexant in a 1:1 ratio. Including re-randomised patients would bias the results due to double counting same patients hence, this study was excluded from the SLR*". [3] The EAG is satisfied with this response.

3.2.1 Details of the included trials

The CS[1] identified studies 301[15] and 303[18] as relevant to the decision problem.

3.2.1.1 Study 301

Study 301[15] was a double-blind RCT which enrolled 930 adult and elderly subjects with insomnia disorder, according to the criteria of DSM-5, unless their insomnia was associated with major comorbidities – especially comorbid neurological, affective or psychiatric disorders (e.g., severe or uncontrolled depression or anxiety, dementia) that could interfere with the study endpoints. Participants were randomly assigned to receive daridorexant 50 mg (N=310) or placebo (N=310) for 12 weeks. A further group was randomly assigned to 25 mg (N=310) but results for that group were not reported in the CS[1] as 25 mg is not regarded as the standard dose. The study involved 75 sites across 10 countries (Australia, Canada, Denmark, Germany, Italy, Poland, Serbia, Spain, Switzerland, and the United States (US)), of which 51 sites in seven countries (Canada, Denmark, Germany, Poland, Spain, Switzerland, and the US) enrolled and randomised subjects.

Treatment comprised of single-blind treatment (placebo matching daridorexant, administered during the placebo run-in and run-out periods) and double-blind treatment (daridorexant, or placebo matching daridorexant, administered from randomisation to end of double-blind treatment period) (Table 3.4).

Table 3.4: Trial drugs in Study 301

Drug	Dose	Frequency of administration	Route of administration	Duration
Daridorexant, film coated tablet	25 mg and 50 mg	One tablet taken orally once daily in the evening	Oral	84 ± 2 days
Placebo matching daridorexant, film coated tablet	-			Single-blind placebo run-in period (13–24 days), treatment period (84 ± 2 days), and single-blind placebo run-out period (7 + 2 days)
Based on Table 8, CS[1] CS = company submission; mg = milligram				

Therapies considered necessary for a subject's well-being and not categorised as prohibited concomitant medications could be used in Study 301.[15] However, initiation of new medication was discouraged, and concomitant medication was preferably not changed during the study. The use of non-sedating antihistamines, opioids/narcotics, centrally acting muscle relaxants with psychotropic effects, and pseudoephedrine was permitted with restrictions. Inhaled or nasal corticosteroids were permitted.

The following concomitant therapies were forbidden during Study 301:[15]

- Treatment with another investigational drug until EOS.
- Study-prohibited central nervous system (CNS)-active medications for five half-lives of the respective drug (but at least 2 weeks) prior to Visit 1 and until 24 hours after EOT.
- Treatment with moderate or strong CYP3A4 inhibitors, or moderate or strong CYP3A4 inducers until 24 hours after EOT.

Cognitive behavioural therapy for insomnia was only allowed if the treatment started at least 1 month prior to Visit 3 (baseline) and the subject agreed to continue this CBT-I throughout the study. Initiation of CBT-I during the study was not allowed.

EAG comment:

- No comparison was made across study arms for the number using allowed CBT-I or other treatment options. This had potential to be a confounder if it differed between arms.
- In the clarification letter the company was asked to provide data on the numbers using CBT-I or other allowed treatment options. The company stated that, "*In study 301, only three randomised subjects (0.3%; 1 subject in each treatment group) were treated with CBTi during the study. Thus, CBTi was not expected to be a confounder in the analyses. Other therapies considered necessary for a subject's well-being was allowed during the study 301; however, the use of these therapies at baseline (study 301) and at start of double-blind treatment (study 303) was balanced across the treatment groups and were not expected to contribute as confounding factors in the efficacy and safety analyses. The study design of daridorexant clinical trial excluded patients with acute or critical pathologies to prohibit the use of non-sedating antihistamines, opioids/narcotics, centrally acting muscle relaxants with psychotropic effects, pseudoephedrine, and inhaled or nasal corticosteroids. Further, randomization of trial population ensured demographic and clinical characteristics of patients balanced confounding factors across the treatment arms"*. [3]In relation to CBT-I the EAG fully accepts the company response. However, for other therapies at baseline, no data are provided to back up the assertion that "*use of these therapies …was balanced across the treatment groups*". [3]

3.2.1.2 Study 303

Study 303[18] was an extension study of Study 301[15] and Study 302.[16] Study 303[18] was primarily a comparative safety study, but it included placebo-controlled subjective efficacy data of relevance to assess long-term maintenance with daridorexant.

3.2.1.2.1 Daridorexant group

Subjects assigned to a daridorexant group in Study 301[15] or Study 302[16] were assigned to the same daridorexant dose (i.e., 10 mg, 25 mg or 50 mg) in Study 303.[18] Therefore, because Study 302[16] did not contain any subjects with 50 mg daridorexant, all subjects in Study 303[18] with a 50 mg daridorexant dose were from the 50 mg arm in Study 301[15] (N=137 after attrition).

3.2.1.2.2 Placebo group

Subjects originally randomised to placebo in studies 301[15] and 302[16] were re-randomised to placebo or 25 mg daridorexant, and those assigned to placebo formed the placebo arm for Study 303[18] (N=128). It is not clear how many of the 128 in the placebo group in Study 303 were originally in Study 301.[15]

EAG comment:

- Although Study 303[18] evaluated placebo, and 10 mg, 25 mg and 50 mg of daridorexant, only the 50 mg daridorexant and placebo groups in Study 303[16] are of relevance to this report and only results pertaining to these two groups will be reported.

Table 3.5 summarises the trial drug administration in Study 303.[18]

Table 3.5: Trial drugs in Study 303

Drug	Dose	Frequency of administration	Route of administration	Duration
Daridorexant, film coated tablet	10 mg, 25 mg and 50 mg	One tablet taken orally once daily in the evening	Oral	280 ± 7 days
Placebo matching daridorexant, film coated tablet	-			Treatment period (280 ± 7 days), and single-blind placebo run-out period (7 + 2 days)
Based on Table 28, CS[1] CS = company submission				

Therapies considered necessary for the subject's well-being and not categorised as prohibited concomitant medications could be used in the study, including coronavirus disease 2019 (COVID-19) vaccines.

The following concomitant therapies were forbidden during the study:

- Treatment with another investigational drug until EOS.
- Study-prohibited CNS-active medications from at least 1 week prior to Visit 1 and until 24 hours after end of trial (EOT).

Treatment with moderate or strong CYP3A4 inhibitors or moderate to strong CYP3A4 inducers from at least 1 week prior to Visit 1 until 24 hours after EOT.

Table 3.6 and

Table 3.7 summarise the methodologies of studies 301[15] and 303.[18]

Table 3.6: Study methodology for Study 301

Study	ID-078A301 (NCT03545191)[15]
Study design	Multi-centre, double-blind, randomised, placebo-controlled, parallel-group
Population	Adult (18-64 years) and elderly (≥65 years) male and female subjects with a diagnosis insomnia disorder as per the DSM-5® criteria and moderate-to-severe insomnia as per ISI© (ISI© ≥15).
Inclusion criteria	• Insomnia disorder according to the DSM-5® criteria. • Self-reported insomnia of at least moderate severity (ISI© score ≥15) at screening. • Sleep disturbance causing clinically significant distress or impairment in social, occupational, educational, academic, behavioural, or other important areas of functioning. • Self-reported insufficient sleep quantity (≥30 minutes to fall asleep, wake time during sleep ≥30 minutes, and sTST ≤6.5 hours during the night) for at least 3 nights per week during at least 3 months prior to the screening visit, and for at least 3 out of 7 nights on the SDQ completed during the placebo run-in period prior to the run-in PSG nights. • Objective sleep quantity parameters assessed on two consecutive PSG nights during the placebo run-in period: mean LPS ≥20 minutes, with neither of the two nights <15 minutes; mean WASO ≥30 minutes, with neither of the two nights <20 minutes; and mean TST <420 minutes. • Subjects were required to sign informed consent prior to any study-mandated procedure.
Exclusion criteria	• Subjects self-reporting daytime napping ≥1 hour per day and ≥3 days per week. • Subjects with BMI <18.5 or >40.0 kg/m2. • Subjects who were pregnant, breastfeeding, or planning to become pregnant. • Subjects with any lifetime history of suicide attempt, sleep-related breathing disorders, periodic limb movement disorder, restless legs syndrome, circadian rhythm disorder, REM behaviour disorder, narcolepsy, or apnoea/hypopnea. • Subjects with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments. • Subjects aged ≥50 years with a Mini Mental State Examination© score <25. • Subjects treated with CNS-active drugs; CBT was allowed if started at least 1 month prior to the run-in PSG nights and intended to be continued throughout the study. • Subjects not able or willing to stop treatment with moderate or strong CYP3A4 inhibitors or inducers within at least 1 week prior to the start of the placebo run-in period.
Intervention(s)	Daridorexant (25 mg and 50 mg)*

Study	ID-078A301 (NCT03545191)[15]
Comparator(s)	Placebo
Reported outcomes specified in the decision problem	The outcomes relevant for the decision problem include: 1. Improvement of night-time symptoms of insomnia (WASO, sWASO, LPS) 2. Changes in sleep architecture and sleep efficiency (LPS, TST, sTST) 3. Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function (TST, sWASO, sLSO) 4. Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score 5. Safety and tolerability (adverse events, next morning residual effect, rebound insomnia, abuse potential, SDS©) 6. HRQoL (ISI© score)
All other reported outcomes	1. Withdrawal symptoms 2. Sleep continuity (WASO by quarter of the night and by hour of the night, TST by quarter of the night, sleep awakenings measured by PSG or self-reported) 3. Sleep efficiency 4. PGA-S, and PGI-C scores

Based on Table 5 and table 7, CS[1]
*Only the evidence for daridorexant 50 mg versus placebo is presented in this submission
BMI = body mass index; CBT = cognitive behavioural therapy; CNS = central nervous system; CS = company submission; CYP3A4 = cytochrome P450 3A4; DSM®-5 = Diagnostic and Statistical Manual of Mental Disorders®, Fifth Edition; HRQoL = health-related quality of life; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index©; LPS = latency to persistent sleep; PGA-S = Patient Global Assessment of Disease Severity; PGI-C = Patient Global Impression of Change; PICO = population, intervention, comparator and outcome; PSG = polysomnography; REM = rapid eye movement; SDQ = Sleep Diary Questionnaire; SDS© = Sheehan Disability Scale©; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; sWASO = subjective wake time after sleep onset; TST = total sleep time; VAS = visual analogue scale; WASO = wake time after sleep onset

Table 3.7: Study methodology for Study 303

Study	ID-078A303 (NCT03679884)[18]
Study design	Multi-centre, double-blind, parallel-group, randomised, placebo-controlled, three doses, 40-week safety extension study to ID-078A301 and ID-078A302
Population	Adult (18-64 years) and elderly (≥65 years) male and female subjects with insomnia disorder according to DSM-5® criteria, who had completed daridorexant treatment in Study 301 and Study 302
Inclusion	• Signed informed consent prior to any study-mandated procedure (Visit 1). • Completion of the double-blind study treatment and placebo run-out period of 301 or 302 (Visit 1). • For woman of childbearing potential, the following was required: • Negative urine pregnancy test (EOT of 301 or 302 studies). • Agreement to use the contraception scheme as required by the protocol from Visit 1 up to at least 30 days after EODBT.
Exclusion	• Subjects self-reporting daytime napping ≥1 hour per day and ≥3 days per week. • Subjects with BMI <18.5 or >40.0 kg/m^2. • Subjects who were pregnant, breastfeeding, or planning to become pregnant. • Subjects with any lifetime history of suicide attempt, sleep-related breathing disorders, periodic limb movement disorder, restless legs syndrome, circadian rhythm disorder, REM behaviour disorder, narcolepsy, or apnoea/hypopnea. • Subjects with acute or unstable psychiatric conditions, suicidal ideation with intent, alcohol or drug abuse, or with history or clinical evidence of any disease, medical condition or treatment that could affect the subject's safety or interfere with the study assessments. • Subjects aged ≥50 years with a Mini Mental State Examination© score <25. • Subjects treated with CNS-active drugs; CBT was allowed if started at least 1 month prior to the run-in PSG nights and intended to be continued throughout the study. • Subjects not able or willing to stop treatment with moderate or strong CYP3A4 inhibitors or inducers within at least 1 week prior to the start of the placebo run-in period.
Intervention(s)	Daridorexant (10 mg, 25 mg and 50 mg)*
Comparator(s)	Placebo
Reported outcomes specified in the decision problem	The outcomes of the decision problem include: 1. Safety and tolerability (adverse events, next morning residual effect, rebound insomnia, abuse potential) 2. Improvement of night-time symptoms of insomnia (sWASO) 3. Changes in sleep architecture and sleep efficiency (sTST) 4. Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function (sLSO) 5. Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score 6. HRQoL (ISI© score)

Study	ID-078A303 (NCT03679884)[18]
All other reported outcomes	1. SDQ VAS 2. Withdrawal symptoms 3. Self-reported awakenings 4. PGA-S and PGI-C scores

Based on Table 6, CS[1]
*Only the evidence for daridorexant 50 mg versus placebo is presented in this submission
CBT = cognitive behavioural therapy; CNS = central nervous system; CS = company submission; CYP3A4 = cytochrome P450 3A4; DSM®-5 = Diagnostic and Statistical Manual of Mental Disorders®, Fifth Edition; HRQoL = health-related quality of life; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index©; PGA-S = Patient Global Assessment of Disease Severity; PGI-C = Patient Global Impression of Change; PICO = population, intervention, comparator and outcome; PSG = polysomnography; SDQ = Sleep Diary Questionnaire; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; sWASO = subjective wake time after sleep onset; TST = total sleep time; VAS = visual analogue scale; WASO = wake time after sleep onset

EAG comment:

- Neither study included participants from the UK, as only Canada, Denmark, Germany, Poland, Spain, Switzerland, and the US enrolled and randomised participants. One percent of participants in Study 301[15] were Asian, 9.5% were Black and 89.5% were White. In Study 303,[18] 1% of participants were Asian, 8.5% were Black and 89.5% were White. This is different to the overall UK population as measured in the 2011 census [19], where 7.5% of the population are Asian, 3.3% of the population are Black and 86% of the population are White (the analogous information from the 2021 census is not currently available). Of course, the ethnic proportions in the *overall* UK population are not necessarily the same as those in the UK population of *chronic insomnia* patients, because there may be an interaction between incidence of chronic insomnia and ethnicity. For example, Fernandez-Mendoza et al. 2021[20] showed that persistence of insomnia in ethnic minorities in the United States of America (USA) is double that in non-Hispanic Whites in a high socio-economic status stratum, strongly suggesting an interaction where non-White ethnicity leads to an increase in the incidence of chronic insomnia. Unfortunately, the ethnic proportions in the UK population of chronic insomnia patients do not appear to be available in the literature, so there is uncertainty whether the ethnic proportions of the trial are representative of the ethnic proportions of the UK target population. Any difference in ethnicity proportions between trial and UK target population could potentially have an impact on applicability, if ethnicity is an outcome modifier; that is, if there is an interaction between ethnicity and the efficacy of daridorexant.
- There is no evidence from the sub-group analyses for Study 303[18] that ethnicity is an outcome modifier, but ethnicity was only evaluated as a sub-grouping variable for sTST and Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) in Study 303. In addition, ethnicity was not evaluated as a sub-grouping variable for any outcome in Study 301.[15] Therefore, doubts about applicability must remain. In the clarification letter the company were asked to comment on the generalisability of the trial population characteristics to the patient population in England and Wales. The company stated that: "*The company acknowledges the difference in ethnic distribution between the trial population and the patient population in England and Wales. However, as highlighted by the EAG, there is no evidence from the subgroup analyses for study 303 that ethnicity is an outcome modifier. Although this was only evaluated for sTST and IDSIQ, the company expects this to be applicable to all other primary and secondary endpoints.*" [3] The EAG response to this is that the company cannot know this until they have evaluated it. Therefore, uncertainty exists, and it is possible that ethnicity is an outcome modifier.
- The company were also asked to provide data sub-grouped for ethnicity for all primary and secondary outcomes in both Study 301 and Study 303. The company responded by stating that, "*The small sample size of the Asian and Black subgroups precluded meaningful comparison of all primary and secondary outcomes across ethnic subgroups. As mentioned in the response to A22 (a), based on the subgroup analysis for study 303, there is no evidence that ethnicity is an outcome modifier for sTST and IDSIQ and the company expects this to be applicable to all other primary and secondary endpoints*". [3] The EAG does not accept this response. The EAG critique is that if the sub-group sample sizes were sufficient for the sTST and IDSIQ outcomes, they would have been sufficient for the other outcomes, where similar sample sizes were observed.
- There was little information provided on comorbid conditions in participants in the trials, which was surprising given that insomnia disorder is associated with various comorbid conditions such as chronic obstructive pulmonary disease, heart failure, chronic pain, and psychiatric conditions (depression, anxiety, substance abuse, and post-traumatic stress disorder). In the clarification letter, the company was asked to provide details on the clinical characteristics of any other pathologies present in the trial populations. The company stated that, "*In study 301, previous psychiatric*

disorders were reported for 54 subjects (5.8%), of which the most common was depression (24 subjects, 2.6%); additionally, major depression was reported for 7 subjects (0.8%) and anxiety for 4 subjects (0.4%). Previous nervous system disorders were reported for 29 subjects (3.1%), of which the most common was migraine (6 subjects, 0.6%). Study-concomitant medical conditions (excluding conditions and symptoms related to insomnia) were reported for 646 subjects (69.5%) and were balanced across the treatment groups. Table 3.8 *illustrates the study concomitant medical conditions by primary system organ class and preferred term in the overall population of study 301."* [3]

Table 3.8: Study concomitant medical conditions by primary system organ class

System Organ Class Preferred Term	Total N=930; n (%)
Psychiatric disorders	43 (4.6%)
Tobacco abuse	15 (1.6%)
Anxiety	8 (0.9%)
Depression	4 (0.4%)
Nervous system disorders	121 (13.0%)
Headache	52 (5.6%)
Migraine	22 (2.4%)
Somnolence	12 (1.3%)
Metabolism and nutrition disorders	234 (25.2%)
Hypercholesterolaemia	74 (8.0%)
Obesity	70 (7.5%)
Type 2 diabetes mellitus	43 (4.6%)
Vascular disorders	223 (24.0%)
Hypertension	207 (22.3%)
Musculoskeletal and connective tissue disorders	198 (21.3%)
Osteoarthritis	73 (7.8%)
Mack pain	38 (4.1%)
Endocrine disorders	87 (9.4%)
Hypothyroidism	72 (7.7%)
Based on Table 7, from clarification question response from company[3]	

3.2.2 Statistical analyses of the 301[15]/303[18] studies

The statistical analyses used for the studies 301[15] and 303[18] are presented in

Table 3.9 and

Table 3.10.

Table 3.9: Summary of statistical methods and analysis sets of Study 301

Study name (number)	Study 301 (NCT03545191)[15]
Research hypothesis relevant to NICE scope	For each of the primary endpoints (change from baseline in WASO [sleep maintenance] and LPS [sleep onset], and secondary endpoints (change from baseline in sTST [sleep quantity], and IDSIQ sleepiness domain score [daytime function], four null hypotheses were defined as follows: H1: Daridorexant 50 mg – Placebo = 0 at Month 1 H2: Daridorexant 50 mg – Placebo = 0 at Month 3 where 'Daridorexant 50 mg', and 'Placebo' represent the mean change from baseline for the given endpoint (WASO, LPS, sTST or IDSIQ sleepiness domain score) and time point (Month 1 or Month 3).
Analysis sets	Screened analysis set: The screened analysis set comprised all subjects who entered screening and received a subject identification number. Full analysis set: The FAS comprised all subjects assigned (i.e., randomised) to a double-blind study treatment. In order to adhere to the intention-to-treat principle as much as possible: Per-protocol set: The per-protocol set comprised all subjects from the FAS who received at least one dose of double-blind study treatment and who complied with the protocol sufficiently to be likely to exhibit the treatment effects. Safety set: The safety set comprised all subjects who received at least one dose of double-blind study treatment. Treatment withdrawal set: The treatment withdrawal set comprised all subjects in the safety set who received at least one dose of single- blind placebo treatment in the placebo run-out period.
Statistical analysis for primary and key secondary efficacy endpoints	Analysis of the primary and secondary efficacy endpoints was performed on the FAS. Linear mixed effects model was used for the analysis of change from baseline in WASO, LPS, sTST and IDSIQ sleepiness domain score, separately. The analysis model adjusted for the baseline value of the relevant response variable (either WASO, LPS, sTST or IDSIQ sleepiness domain score), age group (<65; ≥65 years), treatment (daridorexant 50 mg; placebo), visit (Month 1; Month 3), and the interaction of treatment by visit, and baseline by visit. To evaluate the efficacy hypotheses, appropriate contrasts were computed to test the treatment differences of interest (i.e., the difference in LSM change from baseline between daridorexant and placebo, both at Month 1 and Month 3).
Statistical analysis for other efficacy endpoints	Analysis of the other efficacy endpoints was performed on the FAS. The same model as for the main analysis of the primary and secondary endpoints (linear mixed effects model), was fitted for TST, sWASO, sLSO and IDSIQ scores (total score; alert/cognition and mood domain scores). The LSM for each treatment group was reported with associated SEs and 95% CIs. The placebo-adjusted LSM was displayed with associated SE, 95% CI and unadjusted two-sided p-value. Other efficacy endpoints (change from baseline to Month 1 and Month 3 in TST, sWASO, sLSO, and IDSIQ total, alert/cognition domain, and mood

Study name (number)	Study 301 (NCT03545191)[15]
	domain scores), with their observed values, were summarized descriptively.
Statistical analysis of exploratory endpoints	Analysis of the exploratory efficacy endpoints was performed on the FAS. The exploratory endpoints (change from baseline to Month 1 and Month 3 of the respective variables) were summarised descriptively with the observed values.
Statistical analysis of safety endpoints	All safety endpoints were summarised descriptively.
Sample size & power calculation	The assumptions for the between-subject SD per treatment group for WASO, LPS, and sTST were based on the two phase II studies (201 and 202) conducted in adult and elderly subjects with insomnia receiving 5 mg, 10 mg, 25 mg, 50 mg daridorexant or placebo. The difference compared to placebo in the mean change from baseline to Month 1 and Month 3 was assumed to be 15 (WASO and LPS) and 20 minutes (sTST). Based on a two-sample z-test, at least 900 subjects randomised to 50 mg daridorexant, 25 mg daridorexant, and placebo in a 1:1:1 ratio (i.e., 300 per group) would provide 98.9% power to detect an effect size of 0.37 for a single hypothesis test. This accounts for the Bonferroni correction, where the significance level (alpha) is halved and set to 2.5% two-sided. However, as the number of null hypotheses (endpoints) to test increases, the power decreases. The power calculation assumed all null hypotheses were independent (a conservative assumption for power calculations). Consequently, 900 subjects provided at least 90% power to detect an effect size of 0.37 for testing nine independent null hypotheses.
Data management, patient withdrawals	Handling of partially missing data: Partially missing data for WASO and LPS values were handled as follows: if one of the two values was missing either for baseline, Month 1 or Month 3, the single value available was used as the mean for that time point. If both values were missing for a time point, then the mean value was considered missing for that time point. The same approach was used for the following variables: TST, number of shifts from S2, SWS or REM to S1 or awake, number of awakenings, Coding sub-test©, SDS©, and neurological examination. For sTST and IDSIQ sleepiness domain scores, subjects had to have at least 2 days of data during each week to calculate a weekly mean. Otherwise, the mean value was considered missing for that week. The same approach was used for the following variables: sWASO, sLSO, IDSIQ scores (total score, alert/cognition domain and mood domain scores), VAS scores, and number of self-reported awakenings.

Based on Table 10, CS[1]

CS = company submission; FAS = full analysis set; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; LPS = latency to persistent sleep; LSM = least squares mean; NICE = National Institute for Health and Care Excellence; REM = rapid eye movement; S1 = sleep stage 1; S2 = sleep stage 2; SD = standard deviation; SDS© = Sheehan disability scale©; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; SWS = slow-wave sleep; sWASO = subjective wake time after sleep onset; TST = total sleep time; VAS = visual analogue scale; WASO = wake time after sleep onset

Table 3.10: Summary of statistical methods and analysis sets of Study 303

Study name (number)	Study 303 (NCT03679884) [18]
Analysis sets	Enrolled set: The enrolled set included all subjects who completed Study 301 or Study 302 and who consented to enter Study 303. Full analysis set: The FAS comprised all subjects assigned (i.e., randomised) to a study treatment. Safety set: The safety set comprised all subjects who received at least one dose of double-blind study treatment. Treatment withdrawal set: The treatment withdrawal set comprised all subjects in the safety set who received at least one dose of single-blind placebo treatment in the placebo run-out period.
Statistical analysis of safety endpoints	All safety endpoints were summarised descriptively.
Statistical analysis for exploratory efficacy endpoints	Analysis of exploratory efficacy endpoints was performed using the FAS. Linear mixed effects model was used for the analysis of change from confirmatory baseline in sTST, sWASO, sLSO and IDSIQ total score, sleepiness domain, alert/cognition domain, and mood domain scores, separately. The analysis model adjusted for the confirmatory baseline value of the relevant response variable (either sWASO, sLSO, sTST, or IDSIQ total score, sleepiness domain, alert/cognition domain, or mood domain scores), age group as per assigned strata (<65; ≥65 years), treatment (daridorexant 50 mg; placebo), visit (at Month 6 [Week 12 of extension study]; Month 9 [Week 24]; Month 12 [Week 36]), and the interaction of treatment by visit, and baseline by visit. Appropriate contrasts were used to test the difference in LSM change from confirmatory baseline between daridorexant 50 mg and placebo at Month 6 [Week 12]; Month 9 [Week 24]; and Month 12 [Week 36]. Observed values and change from baseline over time in ISI© were summarised descriptively.
Sample size & power calculation	As Study 303 was an extension of studies 301 and 302, no formal sample size calculation was undertaken. It was expected that approximately 1,260 subjects (i.e., ~70% of the total subjects in studies 301 and 302) would enter the extension study, assuming all sites participated in this study.
Data management, patient withdrawals	Handling of missing data: For sTST, sWASO, sLSO, each IDSIQ domain and total scores, VAS scores and number of self-reported awakenings, at least 2 days of data during each week were required to calculate a weekly mean. Otherwise, the mean value was considered missing for that week. The approach implies implicit imputation: the missing data points were given the same value as the mean of the non-missing data points of that same time point or week.

Based on Table 30, CS[1]

CS = company submission; FAS = full analysis set; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index©; LSM = least squares mean; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; sWASO = subjective wake time after sleep onset; VAS = visual analogue scale

EAG comment:

- Statistical approach in both studies appears to be rigorous and correct in general. However, it was unclear which intention to treat (ITT) analyses were used and for which outcomes. This has been examined in the clarification questions. The company defined their ITT analyses as follows: "*Intention-to-treat population was defined as all participants who were randomly assigned to a double-blind study treatment. In order to adhere to the intention-to-treat principle as much as possible:*
    - *Subjects were evaluated according to the treatment and strata they were assigned to, which may differ from the treatment they received;*
    - *All available data were included.*
    - *Intention-to-treat population was analysed in study 301 for the primary and secondary endpoints which included, objective assessments of WASO and LPS, and subjective assessments of TST and IDSIQ sleepiness domain".* [3]

The EAG is satisfied with this clear response.

- In the power calculation for Study 301, an effect size of 0.37 was chosen as the measure of clinical significance, which is normally regarded as a small to medium effect size. This meant that a large number of participants (N=900) were required to achieve 90% power, and therefore that if such a sample size target were attained, more than enough statistical power would be available to detect larger (and more easily discerned) effect sizes (such as >0.5) that might customarily be regarded as denoting a clinical benefit. As all 900 participants were successfully recruited, this was, of course, as Sellar and Yeatman might say, 'A Good Thing'. However, there was concern in the EAG that application of this power analysis might lead to some between-arm differences being statistically significant without being truly clinically significant (that is, having an effect size as low as 0.37). In the clarification letter the company was asked to justify any clear clinical benefits at the level of 0.37 effect size. The company stated that, "*There is a lack of evidence to support the use of a particular measure or combination of measures to demonstrate a clear clinical benefit for patients presenting with insomnia. Instead, an extensive list of outcomes was presented in the CS to provide a holistic assessment of the efficacy and safety of daridorexant. The company has established a meaningful threshold of 55 minutes for sTST compared to baseline using the dose response curve from a phase 2 study. It is challenging to establish a meaningful threshold compared to placebo since placebo effects are often large in insomnia studies*". [3] The EAG does not think that this response answered the question, as the effect size of 0.37 has not been justified as being clinically significant.

3.2.3 Baseline characteristics of Study 301[15] and Study 303[18]

Table 3.11 and Table 3.12 summarise the baseline characteristics of studies 301[15] and 303.[18]

Table 3.11. Baseline characteristics of subjects in Study 301

Variable Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
Age at screening (years)		
Mean (SD)	55.5 (15.3)	55.1 (15.4)
Median (Min, Max)	58 (21, 86)	58 (19, 83)
Sex [n(%)]		
Male	111 (35.8)	100 (32.3)

Variable Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
Female	199 (64.2)	210 (67.7)
Race [n(%)]		
Black or African American	30 (9.7)	28 (9.0)
American Indian or Alaska Native	1 (0.3)	0
Native Hawaiian or other Pacific Islander	1 (0.3)	0
Asian	4 (1.3)	2 (0.6)
White	274 (88.4)	278 (89.7)
Other	0	2 (0.6)
Ethnicity [n(%)]		
Hispanic or Latino	44 (14.2)	51 (16.5)
Not Hispanic or Latino	265 (85.5)	259 (83.5)
Unknown	1 (0.3)	0
BMI (kg/m^2) at screening		
Mean (SD)	26.273 (4.275)	26.428 (4.118)
Region [n(%)]		
US	97 (31.3)	104 (33.5)
Other (non-US)	213 (68.7)	206 (66.5)
WASO (min)		
n	309	309
Mean (SD)	95.484 (37.813)	102.511 (40.766)
LPS (min)		
n	309	309
Mean (SD)	63.619 (37.389)	66.535 (39.769)
sTST (min)		
n	309	309
Mean (SD)	313.178 (57.597)	315.886 (53.144)
IDSIQ sleepiness domain score		
n	309	308
Mean (SD)	22.479 (7.207)	22.260 (6.947)
ISI© score		
n	308	309
Mean (SD)	19.3 (4.0)	19.2 (4.0)

Based on Table 11, CS[1]
BMI = body mass index; CS = company submission; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; ISI© = Insomnia Severity Index©; LPS = latency to persistent sleep; SD = standard deviation; sTST = subjective total sleep time; US = United States; Higher IDSIQ sleepiness domain score represents greater burden of illness; WASO = wake time after sleep onset

EAG comment:

- In general, the differences in baseline characteristics between arms are small and consistent with the magnitude which would be expected with samples of this size. The comparability is therefore

largely consistent with successful randomisation. However, there is a greater baseline WASO value in the placebo arm, which is larger than that which would be expected from random sampling error (mean difference (MD): -7.02 (95% confidence interval (CI): -13.2 to -0.82)). This could be a type I error, given the number of baseline characteristics observed, and therefore does not threaten the general conclusion that the randomisation was successful, but this between-arm baseline difference might still lead to a spurious benefit for the daridorexant group at follow-up in the analysis for that specific outcome. Even so, the use of change scores in the follow-up analysis should eliminate most of the risk of bias from this small baseline difference, and the EAG is therefore satisfied that this between-arm difference is not a major cause for concern.

Table 3.12. Baseline characteristics of subjects in Study 303

Variable Statistic	Daridorexant 50 mg N=137	Placebo N=128
Age at screening (years)		
Mean (SD)	56.9 (13.6)	59.2 (12.6)
Median (Min, Max)	59 (22, 81)	61 (30, 85)
Sex [n(%)]		
Male	39 (28.5)	36 (28.1)
Female	98 (71.5)	92 (71.9)
Race [n(%)]		
Black or African American	15 (10.9)	8 (6.3)
American Indian or Alaska Native	1 (0.7)	0
Native Hawaiian or other Pacific Islander	0	1 (0.8)
Asian	0	2 (1.6)
White	121 (88.3)	115 (89.8)
Other	0	2 (1.6)
Ethnicity [n(%)]		
Hispanic or Latino	19 (13.9)	10 (7.8)
Not Hispanic or Latino	118 (86.1)	118 (92.2)
BMI (kg/m^2) at screening		
Mean (SD)	25.890 (4.238)	25.904 (4.039)
Region [n(%)]		
US	36 (26.3)	46 (35.9)
Other (non-US)	101 (73.7)	82 (64.1)
sTST (min)		
n	137	128
Mean (SD)	303.792 (65.084)	305.071 (56.506)
IDSIQ sleepiness domain score		
n	137	128
Mean (SD)	22.374 (6.562)	21.792 (6.564)
IDSIQ total score		
n	137	128

Variable Statistic	Daridorexant 50 mg N=137	Placebo N=128
Mean (SD)	74.864 (23.519)	70.297 (22.125)
IDSIQ alert/cognition domain score		
n	137	128
Mean (SD)	32.389 (9.999)	30.826 (9.138)
IDSIQ mood domain score		
n	137	128
Mean (SD)	20.101 (8.014)	17.679 (8.005)
sLSO (min)		
n	137	128
Mean (SD)	63.409 (40.300)	64.821 (39.952)
sWASO (min)		
n	137	128
Mean (SD)	80.114 (57.327)	82.675 (52.388)

Based on Table 32, CS[1]
*All demographic data reported in this table are from the respective confirmatory Study 301[15]
BMI = body mass index; CS = company submission; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; Max = maximum; Min = minimum; min = minutes; SD = standard deviation; sLSO = subjective latency to sleep onset; sTST = subjective total sleep time; sWASO = subjective wake after sleep onset; US = United States

EAG comment:

- In general, the differences in characteristics between arms are small and consistent with the magnitude that would be expected with samples of this size. The comparability observed is therefore consistent with successful randomisation.

3.2.4 Risk of bias assessment of Study 301[15] and Study 303[18]

A RoB assessment of Study 301[15] was provided in Appendix D of the CS using the Cochrane RoB tool. This assigned a rating of low risk to all domains of bias, and therefore a rating of low RoB overall. A RoB assessment for Study 303[18] was not carried out.

EAG comment:

- Based on the information in the CSR, the EAG carried out its own risk of bias appraisal for study 303[18]. This showed that allocation concealment was used, blinding was strictly adhered to and an ITT approach was used to minimise attrition bias. The EAG risk of bias rating was therefore deemed 'low'. In addition the EAG reviewed the CSR report for study 301,[15] which confirmed that the RoB in study 301 was low, for the same reasons.
- No comparison was made across study arms for the number using allowed CBT-I or other treatment options. This had potential to be a confounder if it differed between arms. In the clarification letter the company was asked to provide data on the numbers using CBT-I or other allowed treatment options, and the company response to this has been outlined in Section 3.2.1.

3.2.5 Efficacy results of Study 301[15] and Study 303[18]

The final NICE scope lists the following outcomes that need to be covered in the Technology Assessment (TA):

- Resolution of symptoms
- Changes in sleep patterns and architecture
- Sleep quality
- Daytime alertness
- Recurrence of insomnia
- Adverse effects of treatment (including residual daytime sedation and memory impairment)
- HRQoL

The outcomes looked at by the company are as follows:

- Improvement of night-time symptoms of insomnia
- Changes in sleep architecture and sleep efficiency
- Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function
- Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score
- Rebound insomnia
- HRQoL
- Adverse effects of treatment

The first six of these outcomes will now be evaluated in turn. Although outcomes have been presented at time points prior to the longest available follow-up points in the 301[15] and 303[18] studies, only the results at the longest available follow points in each study (normally 3 months and 12 months respectively) will be presented. Adverse outcomes will be evaluated in Section 3.2.6.

EAG comment:

- The CS[1] includes all of the NICE scope outcomes except two: 1) 'resolution of symptoms', which is replaced by 'improvement in night-time symptoms of insomnia' (see Section 2.4 for discussion of this important issue), and 2) recurrence of insomnia, which is not replaced by any directly analogous outcome, except perhaps for 'rebound insomnia' (which although a subsidiary category of insomnia recurrence, does not fully encompass it). The CS[1] also includes additional outcomes that are highly correlated with some of the NICE outcomes and each other, which increases the probability of detecting differences between arms. For example, for the NICE scope outcome of 'daytime alertness' the company has 'daytime alertness', 'daily ability to function', 'daytime functioning as measured by IDSIQ', 'sleepiness' and 'alert/cognition'. The company has been asked to clarify these outcomes and their priority in the clarification letter, and the company responses have been outlined in Section 2.4.
- It is unclear why some validated and commonly used measurement tools were not used, such as the Pittsburgh Sleep Quality Index (PSQI). In the clarification letter, the company was asked why sleep duration was not objectively measured. The company responded as follows: "*The BAP guidelines recommend comprehensive assessment of subjective symptoms of insomnia disorder; objective measures, such as wearable devices/actigraphy or polysomnography (PSG) are indicated if sleep disorders such as sleep apnoea or narcolepsy are suspected. While wearables/actigraphy makes it convenient for trial subjects to measure sleep measures objectively, it tends to be less accurate than PSG and may not be sufficiently sensitive to detect changes in sleep parameters over time. In*

addition, as actigraphy assesses sleep based on movement, it is less accurate when evaluating fragmented sleep, reduced sleep time and/or restless sleep commonly seen in patients with insomnia disorder. Total sleep time (TST) was assessed both subjectively and objectively. But the objective assessment of TST was an exploratory efficacy endpoint of study 301. TST was defined as the time scored as non-awake from lights off to lights on, as determined by PSG. The results of the objective measure are presented in Section B.2.4.4, Table 15." [3] The EAG is satisfied with this response.

- The company was also asked why sleep onset latency was not measured. The company stated that: "*Sleep onset latency was measured as objective subjective endpoints in study 301. Objectively it was assessed as a primary endpoint, latency to persistent sleep (LPS) and subjectively as an exploratory endpoint, latency to sleep onset (LSO) in studies 301 (Section B.2.4.1 and B.2.4.4). LPS was the time from start of recording to the beginning of the first continuous 20 epochs (i.e., 10 min) scored as non-awake. Subjective LSO was the time reported by the subject in answer to the sleep diary questionnaire "How long did it take you to fall asleep?"*. [3] The EAG is satisfied with this response.
- In the clarification letter the company has been asked to confirm that the latest data cut-off was 22 July 2020 and provide newer data, if available, in an addendum. The company confirmed that it was 22 July 2020.

3.2.5.1 Improvement of night-time symptoms of insomnia

3.2.5.1.1 Study 301[15]

Improvement in symptoms from baseline to 3 months were observed to be greater in the daridorexant group. Table 3.13 to Table 3.20 summarise the results for symptom improvement in terms of quality of sleep, depth of sleep, daytime alertness, ability to function, and night-time awakening.

Table 3.13: Change for patient-reported quality of sleep from baseline to Month 3

Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	VAS quality of sleep (mm)	
n	289	289
Mean (SD)	20.21 (22.15)	13.95 (18.85)

Based on Table 1, Appendix M, CS[21]
CS = company submission; n = number; SD = standard deviation; VAS = visual analogue scale

Table 3.14: Change for patient-reported depth of sleep from baseline to Month 3

Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	VAS depth of sleep (mm)	
n	289	289
Mean (SD)	19.24 (22.35)	12.96 (18.59)

Based on Table 1, Appendix M, CS[21]
CS = company submission; n = number; SD = standard deviation; VAS = visual analogue scale

Table 3.15: Change for patient-reported daytime alertness from baseline to Month 3

Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	VAS daytime alertness (mm)	
n	291	288

Mean (SD)	15.99 (20.61)	12.41 (19.16)
Based on Table 1, Appendix M, CS[21] CS = company submission; n = number; SD = standard deviation; VAS = visual analogue scale		

Table 3.16: Change for patient-reported ability to function from baseline to Month 3

Time point Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	VAS Ability to function (mm)	
n	291	288
Mean (SD)	17.12 (22.03)	12.17 (18.28)
Based on Table 1, Appendix M, CS[21] CS = company submission; n = number; SD = standard deviation; VAS = visual analogue scale		

Table 3.17: Change for PGA-S (daytime and night-time symptoms) from baseline to Month 3 in change score

Statistic	Daridorexant 50 mg; N=310			Placebo; N=310		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
PGA-S (daytime symptoms)						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SD	0.93	1.32	1.48	1.02	1.31	1.37
PGA-S (night-time symptoms)						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SD	0.66	0.87	1.03	0.72	0.87	0.88
Based on Table 2, Appendix M, CS[21] CS = company submission; n = number; PGA-S = Patient Global Assessment of Disease Severity; SD = standard deviation						

Table 3.18: Observed value and change for PGI-C (daytime and night-time symptoms) from baseline to Month 3 in change score

Statistic	Daridorexant 50 mg; N=310			Placebo; N=310		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
PGI-C (daytime symptoms)						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SD	0.8	1.39	1.51	0.73	1.26	1.34
PGI-C (night-time symptoms)						
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SD	0.81	1.41	1.58	0.70	1.26	1.37
Based on Table 2, Appendix M, CS[21]						

CS = company submission; n = number; PGI-C = Patient Global Impression of Change; SD = standard deviation

Table 3.19: Observed change from baseline to Month 3 in mean number of PSG awakenings over the whole night, full analysis set

Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	Observed change from baseline to Month 3 in mean number of PSG awakenings over the whole night, full analysis set	
n	287	283
Mean (SD)	0.99 (5.50)	–0.43 (4.99)
Based on Table 4, Appendix M, CS[21] CS = company submission; n = number, PSG = polysomnography, SD = standard deviation		

Table 3.20: Observed change from baseline to Month 3 in number of self-reported awakenings, Full analysis set

Statistic	Daridorexant 50 mg; N=310	Placebo; N=310
	Observed value and change from baseline to Month 3 in number of self-reported awakenings, Full analysis set	
n	289	289
Mean (SD)	–0.66 (1.14)	–0.47 (1.44)
Based on Table 4, Appendix M, CS[21] CS = company submission; n = number, SD = standard deviation		

EAG comment:

- No between-group analyses were presented by the company for the above 10 outcomes, and so mean differences with 95% CIs (for the mean difference of the 3-month change from baseline in outcome on daridorexant minus the 3-month change from baseline in outcome on placebo) have been calculated by the EAG.
- With the exception of the latter two, the 95% CIs demonstrate CIs that do not cross the null line and that favour daridorexant. However, for the between-arm mean difference of the baseline to 3 months change in the mean number of PSG awakenings, there was a significant effect favouring placebo. For the between-arm mean difference of the baseline to 3 months change in the mean number of self-reported awakenings, there was no clear effect in either direction:
    - VAS quality of sleep: MD (95% CI): [redacted]
    - VAS depth of sleep: MD (95% CI): [redacted]
    - VAS daytime alertness: MD (95% CI): [redacted]
    - VAS ability to function: MD (95% CI): [redacted]
    - PGA-S (daytime symptoms): MD (95% CI): [redacted]
    - PGA-S (night-time symptoms): MD (95% CI): [redacted]

- PGI-C (daytime symptoms): MD (95% CI): [redacted]
- PGI-C (night-time symptoms): MD (95% CI): [redacted]
- Mean number of PSG awakenings over night: MD (95% CI): [redacted]
- Mean number of self-reported awakenings: MD (95% CI): [redacted]

- Non-subjective and subjective evaluations of the time awake time after sleep onset (WASO and sWASO) and LPS are also regarded as a measure of night-time symptoms in the CS[1], so have been placed in this category. All were improved by daridorexant compared to placebo over the 3 months of Study 301[15] (Table 3.21 to Table 3.23).

Table 3.21: Between treatment analysis for change from baseline in WASO (min) to Month 3

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	287	-29.41	2.031	[-33.399, -25.427]	-18.30	2.875	[-23.945, -12.661]
Placebo (N=310)	283	-11.11	2.049	[-15.131, -7.088]	-	-	-

Based on Table 13, CS[1]
CL = confidence limit; CS = company submission; LSM = least squares mean; min = minutes; n = number; SE = standard error; WASO = wake time after sleep onset

Table 3.22: Between treatment analysis for change from baseline in sWASO (min) to Month 3

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	[redacted]						
Placebo (N=310)	[redacted]						

Based on Table 15, CS[1]
CL = confidence limit; CS = company submission; LSM = least squares mean; min = minutes; n = number; SE = standard error; sWASO = subjective wake time after sleep onset

Table 3.23: Between treatment analysis for change from baseline in latency to persistent sleep (LPS) (min) to Month 3 full analysis set

Visit / Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL

Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	287	-34.80	1.689	[-38.118, -31.490]	-11.67	2.383	[-16.348, -6.994]
Placebo (N=310)	283	-23.13	1.697	[-26.464, -19.803]	-	-	-
Based on Table 13, CS[1] CL = confidence limit; CS = company submission; LPS = latency to persistent sleep; LSM = least squares mean; min = minutes; n = number; SE = standard error							

EAG comment:

- For the WASO and sWASO outcomes the company provided a between-arm analysis that demonstrated efficacy at a population level.
- In Appendix M,[21] further analyses were presented for these outcomes broken down by quarters of the night and hours of the night. These have not been included in this report to reduce multiplicity of outcomes.

Subjective total sleep time is also regarded as a measure of night-time symptoms[3], so has been placed in this category. Subjective total sleep time improved by daridorexant over the 3 months of Study 301 compared to placebo (Table 3.24).

Table 3.24: Between treatment analysis for change from baseline in sTST (min) to Month 3

Visit	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Treatment group							
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	289	57.67	3.311	[51.171, 64.168]	19.77	4.661	[10.623, 28.918]
Placebo (N=310)	289	37.90	3.315	[31.393, 44.404]	-	-	-
Based on Table 14, CS[1] CL = confidence limit; CS = company submission; LSM = least squares mean; min = minutes; n = number; SE = standard error; sTST= subjective total sleep time							

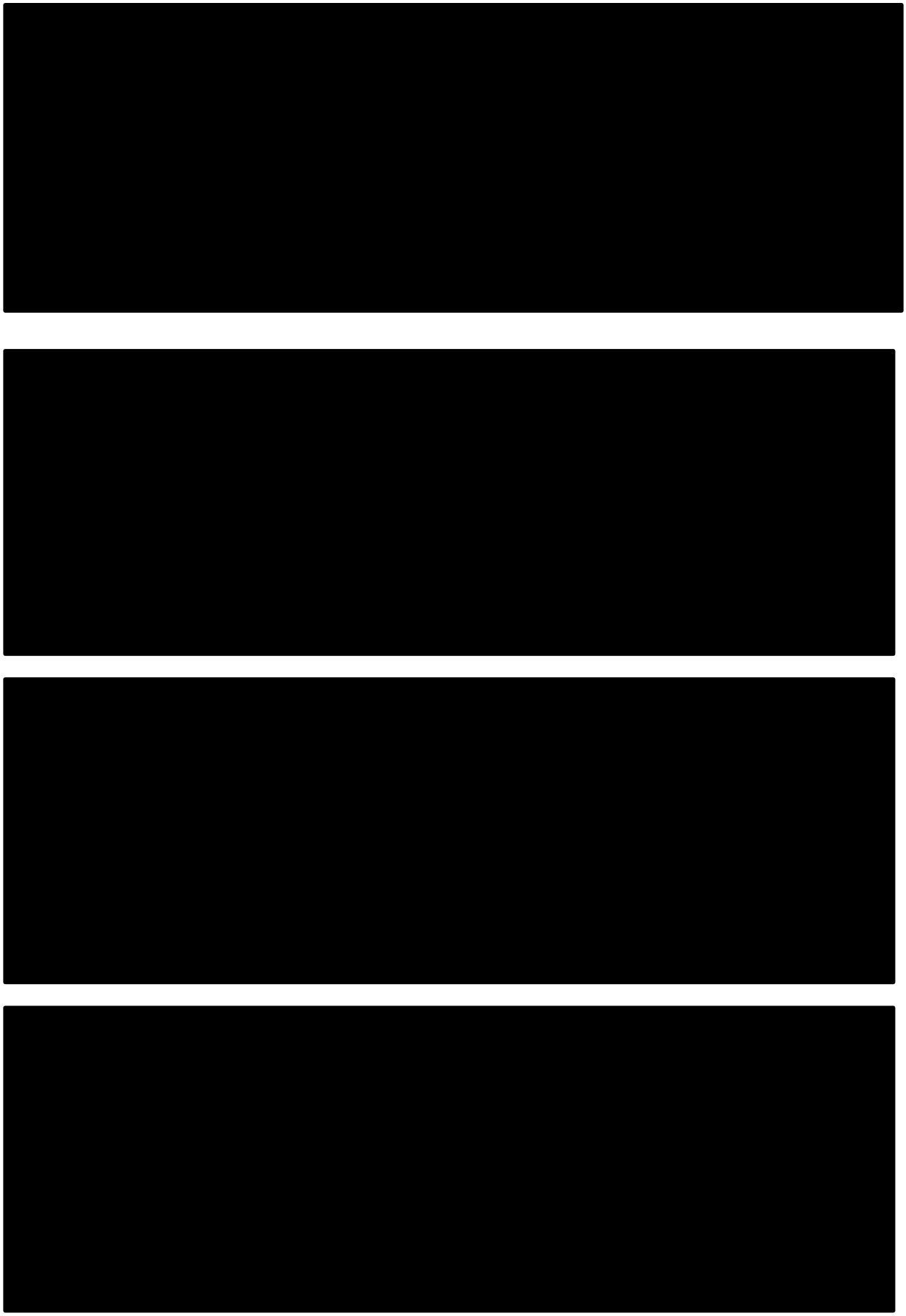
EAG comment:

- The between-arm analysis conducted by the company shows a significant effect for sTST.
- In the Appendix M, a further analysis was presented for sTST broken down by quarters of the night. This has not been included in this report to reduce multiplicity of outcomes.

3.2.5.1.2 Study 303[18]

Patient global assessments of disease severity for daytime symptoms (PGA-S and PGI-C) demonstrated [redacted] in daridorexant 50 mg group compared with placebo (change from the baseline [redacted] for PGA-S [redacted] for PGI-C at [redacted]). Change in mean baseline values of quality of sleep, depth of sleep, daytime alertness, and daily ability to function (assessed on VAS rating scale) for daridorexant 50 mg were numerically greater than placebo at all visits up to Week 40 (Figure 3.1).

Figure 3.1: Observed value and change from baseline over time from baseline to 40 weeks in patient-reported symptoms



Based on Figure 3, Appendix M CS[21]
CS = company submission; VAS = visual analogue scale

EAG comment:

- No between-arm analyses were presented in the CS or appendices. It was not possible for the EAG to carry out a precise between-arm-analysis because data on the number in each arm for this analysis was not provided. However, using the largest n values possible (the n for the full dataset in Study 303), the following MDs (95% CIs) were calculated:
    - PGA-S daytime: MD (95% CI): [redacted]
    - PGI-C daytime: MD (95% CI): [redacted]
- It can be seen that for PGI-C daytime, the 95% confidence intervals of the small MD of [redacted] cross the null line, indicating a probability of >0.05 that the population MD may not be in the same direction of effect as the point estimate. Importantly, this was observed even when using the largest n value available, and therefore the p value would have continued to be at >0.05 at any other possible (necessarily smaller) n value.
- For the results for symptom improvement in terms of quality of sleep, depth of sleep, daytime alertness, and ability to function, no numerical data are given, and the only information is provided in figures (Figure 3.1). It is therefore difficult to differentiate between true population differences and random sample differences.

Subjective wake time after sleep onset was not significantly different between arms at 12 months (Table 3.25), as *shown by the confidence intervals of the between-arm value crossing the null* line.

Table 3.25: Change in baseline to month 12 for sWASO

Visit Treatment group	n	LSM (95% CL)	Difference to placebo	
			LSM (95% CL)	p-value (two-sided)
Between treatment analysis for change from baseline in sWASO (min) to Month 12, full analysis set				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 35, CS[1]
CL = confidence limit; CS = company submission; LSM = least squares mean; min = minutes; n = number; sWASO = subjective wake time after sleep onset

At 12 months a numerical difference between the arms remained for sTST, but this was no longer significant (Table 3.26).

Table 3.26: Between treatment analysis for change from baseline in sTST (min) to Month 12

Visit Treatment group	n	LSM (95% CL)	Difference to placebo LSM (95% CL)
Change from baseline to Month 12			
Daridorexant 50 mg (N=137)	87	[redacted]	[redacted]
Placebo (N=128)	70	[redacted]	[redacted]
Based on Table 33, CS[1] CL = confidence limit; CS = company submission; LSM=least squares mean; min = minutes; n = number; sTST=subjective total sleep time			

3.2.5.2 Changes in sleep architecture and sleep efficiency

3.2.5.2.1 Study 301[15]

Numerically, the duration from LPS to the first epoch of rapid eye movement (REM) sleep, and the latency from sleep onset to the first epoch of REM sleep was [redacted] in participants with daridorexant 50 mg than placebo at Month 3 (Table 3.27). Similarly, the latency from sleep onset to the first epoch of REM was [redacted] on daridorexant 50 mg than on placebo at Month 3 (Table 3.28).

Table 3.27: Sleep architecture: Change from baseline to Month 3 in latency (min) from LPS to REM

Time point statistic	Daridorexant 50 mg; [redacted]	Placebo; [redacted]
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Based on Table 3, Appendix M, CS[21] CS = company submission; LPS = latency to persistent sleep; min = minutes; n = number; REM = rapid eye movement; SD = standard deviation		

Table 3.28: Sleep architecture: Observed values at baseline and Month 3 in latency (min) from sleep onset to REM

Time point statistic	Daridorexant 50 mg; [redacted]	Placebo; [redacted]
Baseline		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Month 3		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
Based on Table 3, Appendix M, CS[21] CS = company submission; min = minutes; n = number; REM = rapid eye movement; SD = standard deviation		

EAG comment:

- Between-arm analyses were not conducted by the company, so the EAG has carried these out below. The MD (95% CIs) for the two outcomes are as follows:
  - Change from baseline to Month 3 in latency LPS to REM [MD [redacted]
  - Latency sleep onset to REM [MD (95%)] at Month 3: [redacted]
- These denote significant differences between arms.
- However, for latency of sleep onset to REM, the company does not present a change score, and the final 3-month analysis that is presented may have been biased by the 6-point difference favouring daridorexant that was already present at baseline. Therefore, the unbiased between-arm difference in efficacy is unclear for this outcome.

Numerically, larger increases were observed from baseline for the participants on daridorexant 50 mg than on placebo in sleep efficiency (Table 3.29).

Table 3.29: Sleep efficiency (%): Change from baseline to Month 3, full analysis set sleep onset latency

Time point statistic	Daridorexant 50 mg N=310	Placebo N=310
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]

Based on Table 8, Appendix M, CS[21]
CS = company submission; mg = milligrams; n = number; SD = standard deviation; % = percentage

EAG comment:

- A between-arm analysis was not conducted by the company, so the EAG has carried this out as follows: the MD (95% CIs) is [redacted].

Numerically, [redacted] from baseline in sleep onset latency (duration from lights off to the first epoch (i.e., 30 seconds) of sleep stage 2 (S2), slow wave sleep (SWS), or REM, or the first 3 consecutive epochs (i.e., 1.5 minutes) of sleep stage 1 (S1)) were observed in participants on daridorexant 50 mg than on placebo, however no statistical comparisons were done (Table 3.30).

Table 3.30: Change from baseline to Month 3 in sleep onset latency (min), full analysis set

Time point Statistic	Daridorexant 50 mg; [redacted]			Placebo; [redacted]		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 8, Appendix M, CS[21]
CS = company submission; mg = milligrams; min = minutes; n = number, SD = standard deviation

EAG comment:

- As a between-arm analysis was not conducted by the company, the EAG has carried this out as follows: the mean difference (MD) (95% CIs) is [redacted]

Total sleep time (TST) is also regarded as measures of sleep architecture and efficiency by the CS[1], so has been placed in this category. Total sleep time improved by daridorexant over the 3 months of Study 301 compared to placebo (Table 3.31).

Table 3.31: Between treatment analysis for change from baseline in TST (min) to Month 3

Visit Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	[redacted]						
Placebo (N=310)	[redacted]						
Between treatment analysis for change from baseline in sWASO (min) to Month 1 and Month 3							
Based on Table 15, CS[1] CS = company submission; CL = confidence limit; LSM = least squares mean; min = minutes; n = number; SE = standard error; sWASO = subjective wake time after sleep onset; TST = total sleep time							

EAG comment:

- The between-arm analysis conducted by the company shows a significant effect for TST.

3.2.5.2.2. Study 303[18]

No data were provided.

3.2.5.3 Changes in quality of sleep, depth of sleep, daytime alertness and daily ability to function

3.2.5.3.1 Study 301[15]

Subjective latency to sleep onset (sLSO) is regarded as a measure of a change in quality of sleep. Daridorexant was observed to lead to [redacted] in sLSO after 3 months compared to placebo (Table 3.32).

Table 3.32: Between treatment analysis for change from baseline in sLSO (min) to Month 3

Visit Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	[redacted]						
Placebo (N=310)	[redacted]						
Based on Table 15, CS[1]							

Visit Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
CL = confidence limit; CS = company submission; LSM = least squares mean; mg = milligrams; min = minutes; n = number; SE = standard error; sLSO = subjective latency to sleep onset							

3.2.5.3.2 Study 303[18]

At 12 months, the difference in improvement between the arms was [redacted] (Table 3.33).

Table 3.33: Change in baseline to month 12 for sLSO (min)

Visit Treatment group	n	LSM (95% CL)	Difference to placebo	
			LSM (95% CL)	p-value (two-sided)
Between treatment analysis for change from baseline in sLSO (min) to Month 12, full analysis set				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 35, CS[1] CL = confidence limit; CS = company submission; LSM = least squares mean; min = minutes; n = number; sLSO = subjective latency to sleep onset;				

3.2.5.4 Daytime functioning as measured by IDSIQ total score, sleepiness, alert/cognition and mood domain score

3.2.5.4.1 Study 301[15]

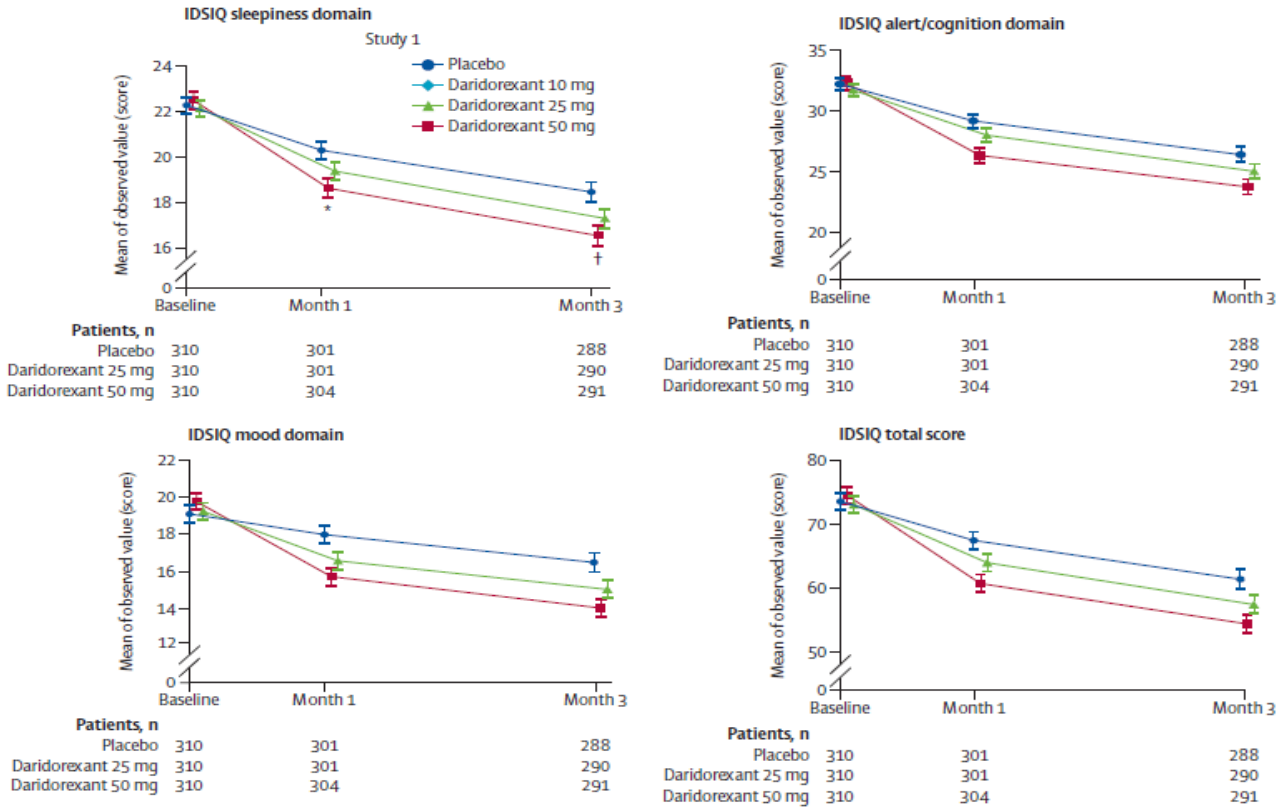
Subjects in the daridorexant 50 mg group reported significant reduction from baseline in IDSIQ sleepiness domain score compared to placebo at Month 3 (least squares mean (LSM) difference -1.90, [-2.95 to -0.98], p=0.0002) (Table 3.34).

Table 3.34: Between treatment analysis for change from baseline in IDSIQ sleepiness domain score to Month 3

Visit Treatment group	n	LSM	SE	95% CL	Difference to placebo		
					LSM	SE	95% CL
Change from baseline to Month 3							
Daridorexant 50 mg (N=310)	291	-5.70	0.361	[-6.405, -4.987]	-1.90	0.510	[-2.905, -0.905]
Placebo (N=310)	288	-3.79	0.363	[-4.503, -3.080]	-	-	-
Based on Table 14, CS[1] CL = confidence limit; CS = company submission; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; LSM = least squares mean; n = number; SE = standard error							

Figure 3.2 summarises the results in the three IDSIQ domains and the total domain at baseline to 3 Months. Differences between daridorexant and placebo for the total score, altert/cognition and mood domain were reported to be p≤0.001 at 3 months.

Figure 3.2: IDSIQ sleepiness domain, IDSIQ alert/cognition domain, IDSIQ mood domain and IDSIQ total score from baseline to 3 Months



Based on Figure 10, CS[1]
Two-sided p-values shown are versus placebo, calculated using the linear mixed effects model for repeated measures. p values for the mood domain, alert/cognition domain, and total score comparisons versus placebo (not adjusted for multiplicity).
CS = company submission; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire

3.2.5.4.2 Study 303[18]

At 12 months, the [redacted] for the daridorexant arm over placebo for the total score and each of the three domains of the ISDIQ persisted (Table 3.35).

Table 3.35: IDSIQ sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score from baseline to month 12

Visit Treatment group	n	LSM (95% CL)	Difference to placebo LSM (95% CL)	Difference to placebo p-value (two-sided)
Between treatment analysis for change from baseline in IDSIQ sleepiness domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]

Visit Treatment group	n	LSM (95% CL)	Difference to placebo LSM (95% CL)	Difference to placebo p-value (two-sided)
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ total score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ alert/cognition domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ mood domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 34, CS[1]
Higher IDSIQ score represents greater burden of illness.

Month 6 timepoint includes the duration of the confirmatory study and corresponds to Week 12 of the extension study, same for Month 9 (Week 24) and Month 12 (Week 36).

Mixed effects model for Repeated Measures: Change from baseline in IDSIQ Sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score = baseline IDSIQ Sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score + stratified age group (<65; >=65 years) + treatment + visit + treatment * visit + baseline * visit

CL = confidence limit; CS = company submission; IDSIQ = Insomnia Daytime Symptoms and Impacts Questionnaire; LSM = least squares mean

3.2.5.5 Rebound insomnia

3.2.5.5.1 Study 301[15]

Recurrence of insomnia (NICE final scope) was not directly assessed in the trial subjects who experienced a treatment effect but those who subsequently discontinued treatment. According to the company, the evidence suggests there was no signal for rebound insomnia after treatment discontinuation (Table 3.36).

Table 3.36 Rebound insomnia, treatment withdrawal set

Time point statistic	Daridorexant 50 mg; N=286			Placebo; N=280		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
Observed value and change from baseline of WASO (min) to run-out treatment withdrawal set						
Run-out - Visit 9						

Time point statistic	Daridorexant 50 mg; N=286			Placebo; N=280		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
n	283	283	283	279	279	279
Mean (SD)	94.743 (37.805)	92.226 (57.394)	-2.517 (52.355)	103.478 (40.708)	83.086 (45.369)	-20.392 (45.776)
Observed value and change from baseline of LPS (min) to run-out treatment withdrawal set						
Run-out - Visit 9						
n	284	284	284	279	279	279
Mean (SD)	63.225 (35.172)	48.190 (49.571)	-15.035 (49.571)	67.829 (40.845)	40.009 (38.390)	-27.820 (47.199)
Observed value and change from baseline of sTST (min) to run-out treatment withdrawal set						
Run-out - Visit 9						
n	281	281	281	274	274	274
Mean (SD)	313.949 (57.920)	356.893 (73.461)	42.943 (59.595)	317.063 (52.238)	359.379 (68.624)	42.316 (52.705)

Based on Table 6, Appendix F, CS[22]
CS = company submission; LPS = latency to persistent sleep; min = minutes; n = number; SD = standard deviation; sTST = subjective total sleep time; WASO = wake time after sleep onset

EAG comment:

- The company did not carry out between-arm statistical analyses. The EAG carried out analyses, showing that the daridorexant versus placebo MD (95% CI) for the change scores were as follows:
  - WASO: [redacted]
  - LPS: [redacted]
- sTST: [redacted]For WASO and LPS these results demonstrate a significantly lower rebound effect for daridorexant than placebo, though this was not observed for sTST.

3.2.5.5.2 Study 303[18]

According to the company, the empirical cumulative distributions by baseline type of change from baseline in sTST from baseline to placebo run out showed [redacted] in any treatment group (Table 3.37).

Table 3.37: Rebound insomnia, treatment withdrawal set

Time point statistic	Daridorexant 50 mg N=93	Placebo N=78
Change from baseline to run out of sTST (min)		
n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]

Based on Table 6, Appendix F, CS[22]
CS = company submission; min = minutes; n = number; SD = standard deviation, sTST = subjective total sleep time

EAG comment:

- The company did not carry out between-arm statistical analyses. The EAG carried out analyses, showing that the daridorexant versus placebo MD (95% CI) for the change scores were as follows:

- sTST: [redacted]
- This confirms [redacted] between arms.

3.2.5.6 Health-related quality of life

The ISI© score was used as the QoL measure for this study.

3.2.5.6.1 Study 301[15]

There was a numerically greater improvement in ISI© score in the daridorexant arm over the 3 months of Study 301[15] (Table 3.38).

Table 3.38: Exploratory endpoint – ISI© score

Time point statistic	n	Mean (SD)
Change from baseline to Month 3		
Daridorexant 50 mg (N=310)	283	–7.2 (6.5)
Placebo (N=310)	281	–5.4 (5.7)
Based on Table 16, CS[1] Change values were calculated only for subjects who had a baseline value. A decrease in score represents an improvement. CS = company submission; ISI© = Insomnia Severity Index©; mg = milligrams; n = number		

EAG comment:

- No between-arm analysis was carried out by the company. A between-arm analysis carried out by the EAG showed that the difference between arms was significant: MD: -1.8 (95% CI: -2.74 to -0.85)

An additional analysis carried out by the company demonstrated a greater proportion in the daridorexant group with a 6 or greater decrement in ISI© score over the first 3 months (Table 3.39). As a decrease in score is an improvement, this represents a benefit for daridorexant.

Table 3.39: Exploratory endpoint – subjects with ≥6 points decrease in ISI© score from baseline to Month 3

	Daridorexant 50 mg ‘ (N=310) n/Nn (%)	Placebo (N=310) n/Nn (%)
Month 3 – 2nd Night	160/283 (56.5)	131/281 (46.6)
Based on Table 17, CS[1] Nn is the number of subjects with non-missing values at the given scheduled visit. CS = company submission; ISI© = Insomnia Severity Index©; n= number		

EAG comment:

- No between-arm analysis was carried out by the company. A between-arm analysis carried out by the EAG showed that the risk ratio (RR) between arms was significant: RR: 1.21 (95% CI: 1.03 to 1.42). This analysis was superfluous, given the previous analysis, and represents over-analysis of data.

3.2.5.6.2 Study 303[18]

There was a [redacted] in ISI© score in the daridorexant arm over the 12 months of Study 303[18] (Table 3.40).

Table 3.40: ISI© score changes from baseline to 40 weeks

Time point statistic	Daridorexant 50 mg (N=137)			Placebo (N=128)		
	Baseline	Post-baseline	Change	Baseline	Post-baseline	Change
n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 36, CS[1] CS = company submission; n = number; SD = standard deviation; ISI© = Insomnia Severity Index						

EAG comment:

- No between-arm analysis was carried out by the company. A between-arm analysis carried out by the EAG showed that the difference between arms was [redacted]

An additional analysis carried out by the company demonstrated a greater proportion in the daridorexant group with a 6 or greater decrement in ISI© score over 12 months (Table 3.41). As a decrease in score is an improvement, this represents a [redacted] for daridorexant.

Table 3.41: Exploratory endpoint – subjects with ≥6 points decrease in ISI© score from baseline to 40 weeks

Time point statistic	Daridorexant 50 mg; N=137 n/Nn (%)	Placebo; N=128 n/Nn (%)
Week 40	[redacted]	[redacted]
Based on Table 37, CS[1] Nn is the number of subjects with non-missing values at the given scheduled visit. CS = company submission; n = number; ISI© = Insomnia Severity Index		

EAG comment:

- No between-arm analysis was carried out by the company. A between-arm analysis carried out by the EAG showed that the relative risk (RR) between arms was [redacted]. Yet again, this analysis was superfluous, given the previous analysis of ISI© at 40 weeks, and represents over-analysis of data.

3.2.5.7 Subgroup analyses

3.2.5.7.1 Study 301[15]

Subgroup analysis was performed in the outcomes of WASO, LPS, sTST and IDSIQ to evaluate the consistency of treatment effect across the following demographic subgroups:

- Age: <65, ≥65 years
- Sex: male, female
- Region: US, other (non-US)

The effect of daridorexant 50 mg on the primary and key secondary efficacy endpoints was reported by the company to be consistent in adults and elderly and across sex and geographical location.

EAG comment:

- The data in Appendix E support the company's statements in the CS[1]: there is little evidence of any appreciable or important effect from age, sex or region on any of the four sub-grouped outcome measures at 3 months. However, it is unknown if there would have been sub-group differences in other outcomes that were not evaluated in this way.

3.2.5.7.2 Study 303[18]

Subgroup analyses of sTST, sWASO, sLSO, IDSIQ domain and total scores were performed to investigate the consistency of the treatment effect across the following subgroups:

Age at screening of confirmatory study: <65, ≥65 years and <75, ≥75 years.

Additionally, the following subgroup analyses were performed for sTST and IDSIQ domain and total scores:

- Sex: Male, female
- Region: US, other (non-US)
- BMI at screening of confirmatory study: <30, ≥30 kg/m^2
- Race: White, Black or African American

[redacted] with the subgroup analysis performed in confirmatory Study 301, there were [redacted] in treatment effect across all subgroups as shown by the [redacted]. Overall, the [redacted] with that of the overall population in extension Study 303[18].

EAG comment:

- The data in Appendix E do not support the company assertion of no sub-group effects for the four outcomes at 12 months. For IDSIQ, the lower body mass index (BMI) sub-group experienced a better response from daridorexant (relative to placebo) than the higher BMI sub-group. For sTST, sWASO, and sLSO there did appear to be some small differences across age sub-groups in terms of the efficacy of daridorexant (relative to placebo) at 12 months, as follows:
    - sTST: >65 versus <65 no difference of any real significance, BUT <75 showed better efficacy for daridorexant versus placebo than <75
    - sWASO: >65 showed better efficacy for daridorexant versus placebo than <65, BUT >75 versus <75 no difference of any real significance
    - sLSO: <65 showed better efficacy for daridorexant vs placebo than >65, AND <75 showed better efficacy for daridorexant versus placebo than <75.
- These effects were highly uncertain, due to the spread of the CIs, but given the likelihood that the study was not powered to detect subgroup differences, and that such sub-group differences may have important implications for applicability, it is important to draw attention to these trends.
- As per the final scope by NICE, "availability and cost of biosimilar and generic products should be considered". These therefore appear to be variables that should be considered in sub-group analyses for both studies 301 and 303 but were not. In the clarification letter, the company was asked to provide a rationale for not including/performing subgroup analyses on these variables. The company responded by stating that, "*Pharmacotherapy is not recommended for long-term management of insomnia disorder. Most of the recommended short-term drugs for insomnia*

disorder are available as generic products. These are not considered as comparators of daridorexant, per the scoping and DPM discussions. Consequently, these analyses were not included in the CS". [3]

3.2.6 Adverse events of Study 301[15] and Study 303[18]

3.2.6.1 Treatment emergent adverse events

3.2.6.1.1 Study 301[15]

During the double-blind study period (0-12 weeks), 37.7% and 34.0% of subjects reported treatment emergent adverse events (TEAEs) in the daridorexant 50 mg group, and placebo group, respectively. Most of the events were reported by the CS[1] to be of mild or moderate intensity (Table 3.42).

Table 3.42: TEAEs during the double-blind study period (12 weeks) reported for ≥2% in any treatment group

Treatment-emergent adverse event	Daridorexant 50 mg N=308; n (%)	Placebo N=309; n (%)
Subjects with at least one event*	116 (37.7)	105 (34.0)
Nasopharyngitis	20 (6.5)	20 (6.5)
Headache	19 (6.2)	12 (3.9)
Accidental overdose	8 (2.6)	5 (1.6)
Fatigue	7 (2.3)	2 (0.6)
Dizziness	7 (2.3)	2 (0.6)
Nausea	7 (2.3)	3 (1.0)

Based on Table 18, CS[1]
*Total number of subjects per treatment group with at least one event. Table is truncated to show only those AEs reported for at least 2% in any treatment group. Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Includes TEAEs occurring (i.e., that started or worsened) during the double-blind study period.
AEs = adverse events; CS = company submission; TEAEs = treatment emergent adverse events

3.2.6.1.2 Study 303[18]

During the double-blind study period, 38.0% and 33.6% of subjects reported TEAEs in the daridorexant 50 mg and placebo groups, respectively (Table 3.43).

Table 3.43: TEAEs reported for ≥2% in any treatment group

Treatment-emergent adverse event	Daridorexant 50 mg N=137; n (%)	Placebo; N=128; n (%)
Subjects with at least one event**	52 (38.0)	43 (33.6)
Nasopharyngitis	11 (8.0)	6 (4.7)
Accidental overdose	4 (2.9)	0
Somnolence	4 (2.9)	0
Fall	3 (2.2)	2 (1.6)
Headache	3 (2.2)	2 (1.6)
Cough	3 (2.2)	0

Treatment-emergent adverse event	Daridorexant 50 mg N=137; n (%)	Placebo; N=128; n (%)
Pneumonia	3 (2.2)	0

Based on Table 38, CS[1]
[*] Includes only those TEAEs occurring (i.e., that started or worsened) during the double-blind study period.
[**] Total number of subjects per treatment group with at least one event. Table is truncated to show only those AE PTs reported for at least 2% in any treatment group.
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA version 22.1
AE = adverse event; MedDRA = Medical Dictionary for Regulatory Activities; PT = preferred terms; TEAE = treatment-emergent adverse event

3.2.6.2 Treatment-emergent serious adverse events in first 12 weeks

3.2.6.2.1 Study 301[15]

The incidence of treatment-emergent serious adverse events (SAEs) was reported in 10 subjects: three (1.0%) and seven (2.3%) subjects in the daridorexant 50 mg and placebo group, respectively (Table 3.44).

Table 3.44: Treatment-emergent SAEs reported at least once in either treatment group

Treatment-emergent SAE	Daridorexant 50 mg N=308; n (%)	Placebo N=309; n (%)
Subjects with at least one event	3 (1.0)	7 (2.3)
Syncope	1 (0.3)[a]	2 (0.6)
Adenocarcinoma of colon	1 (0.3)	0
Haemoglobin decreased	1 (0.3)[a]	0
Post procedural haemorrhage	1 (0.3)[a]	0
Renal colic	1 (0.3)[*]	0
Depression	0	2 (0.6)[b,*]
Anal abscess	0	1 (0.3)
Ankle fracture	0	1 (0.3)
Herpes zoster	0	1 (0.3)
Panic attack	0	1 (0.3)[b]

Based on Table 19, CS[1]
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA dictionary version 22.1. Includes all SAEs occurring from start of double-blind study treatment up to 30 days after the end of double-blind study treatment or enrolment in the ID-078A303 extension study.
a Syncope, haemoglobin decreased, and post procedural haemorrhage were all reported for one subject.
b Depression and panic attack were both reported in the same subject.
*Renal colic and one of the two SAEs of depression occurred during the safety follow-up period.
CS = company submission; SAE = serious adverse event; MedDRA = Medical Dictionary for Regulatory Activities

3.2.6.2.2 Study 303[18]

The incidence of treatment-emergent SAEs was low (5.1% subjects in the daridorexant 50 mg group versus 1.6% subjects in the placebo group) (Table 3.45).

Table 3.45: Treatment-emergent serious adverse events reported at least once in either treatment group

Treatment-emergent SAE	Daridorexant 50 mg N=137; n (%)	Placebo N=128; n (%)
Subjects with at least one event	7 (5.1)	2 (1.6)
Diverticulitis	1 (0.7)	0
Confusional state	1 (0.7)	0
Bone disorder	1 (0.7)	0
Chronic lymphocytic leukaemia	1 (0.7)	0
Influenza like illness	1 (0.7)	0
Pneumonia	1 (0.7)	0
Thyroiditis subacute	1 (0.7)	0
Wrist fracture	1 (0.7)	0
Depression	0	1 (0.8)
Head injury	0	1 (0.8)
Subdural haematoma	0	1 (0.8)
Suicidal ideation	0	1 (0.8)

Based on Table 39, CS[1]
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Preferred terms are based on MedDRA version 22.1 Includes all AEs in the double-blind study period and up to 30 days after double-blind study treatment end date.
AE=Adverse event; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; SAE=Serious adverse event.

3.2.6.3 Adverse events leading to premature discontinuation of double-blind study treatment

3.2.6.3.1 Study 301[15]

The AEs leading to premature study treatment discontinuation were reported for three (1.0%) and 10 subjects (3.2%) in the daridorexant 50 mg, and placebo groups, respectively.

3.2.6.3.2 Study 303[18]

The AEs leading to premature study treatment discontinuation were reported for [redacted] in the daridorexant 50 mg and placebo groups, respectively

3.2.6.4 Adverse events of special interest (AESIs)

3.2.6.4.1 Study 301[15]

AESIs were reported for 3 subjects (2 in daridorexant 50 mg], 1 in placebo). All AESIs were adjudicated as potentially related to study treatment by the ISB (Table 3.46). 'Narcolepsy-like symptoms related to excessive daytime sleepiness' were equally distributed across both treatment groups (1 subject each in the daridorexant 50 mg and placebo groups).'Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations and sleep paralyses' were reported for 1 subject in the daridorexant 50 mg group and none in the placebo group.

All adjudicated AESIs were non-serious, and the majority were of mild intensity, except for 2 events of moderate somnolence and 1 event of severe sleep paralysis. None of the events required treatment, and study treatment continued in all but 1 subject.

Table 3.46: Treatment-emergent AESI after ISB adjudication

Adverse event of special interest	Daridorexant 50 mg N=308; n (%)	Placebo N=309; n (%)
Subjects with at least one event	2 (0.6)	1 (0.3)
Narcolepsy-like symptoms related to excessive daytime sleepiness	1 (0.3)	1 (0.3)
Somnolence	1 (0.3)	1 (0.3)
Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations/sleep paralysis	1 (0.3)	0
Sleep paralysis	1 (0.3)	0

Based on Table 20, CS[1]
Percentages are based on the treatment group N; n = number of subjects with at least one row event. Includes all AESIs, as confirmed by ISB adjudication, occurring from start of double-blind study treatment up to 30 days after the end of double-blind study treatment or enrolment in the ID-078A303 extension study.
AESI = adverse event of special interest; CS = company submission; ISB = Independent Safety Board

3.2.6.4.2 Study 303[18]

Incidence of treatment-emergent adverse event of special interest (AESI) was low, with AESIs reported for two subjects (one in daridorexant 50 mg, one in placebo groups). All AESIs were adjudicated as potentially related to study treatment by the ISB (Table 3.47).

Table 3.47: Treatment-emergent AESIs after ISB adjudication

Adverse event of special interest	Daridorexant 50 mg N=137; n (%)	Placebo N=128; n (%)
Subjects with at least one event	[redacted]	[redacted]
Narcolepsy-like symptoms related to complex sleep behaviour including hallucinations/sleep paralysis	[redacted]	[redacted]
Abnormal dreams	[redacted]	[redacted]
Suicide/self-injury	[redacted]	[redacted]
Suicidal ideation	[redacted]	[redacted]

Based on Table 30, CS[1]
Percentages are based on the treatment group N; n = number of subjects with at least one row event; Subjects may be counted in more than one row. Includes all AEs in the double-blind study period and up to 30 days after double-blind study treatment end date.
AE = adverse event; AESI = adverse event of special interest; CS = company submission; ISB = Independent Safety Board

3.2.6.5 Adverse event sub-groups

3.2.6.5.1 Study 301[15]

The effect of daridorexant on the TEAEs was found to be consistent across all subgroups (Table 3.48).

Table 3.48: Subjects with at least one TEAE during the DB study period by subgroup

	Daridorexant 50 mg; n/N (%)	Placebo; n/N (%)
Overall study population	116/308 (37.7)	105/309 (34.0)
Age at screening (years)		
<65	74/189 (39.2)	67/187 (35.8)
≥ 65	42/119 (35.3)	38/122 (31.1)
<75	108/290 (37.2)	99/292 (33.9)
≥ 75	8/18 (44.4)	6/17 (35.3)
Sex groups		
Male	31/110 (28.2)	30/100 (30.0)
Female	85/198 (42.9)	75/209 (35.9)
BMI at screening (kg/m^2)		
25	52/126 (41.3)	44/117 (37.6)
25-30	44/127 (34.6)	43/135 (31.9)
>30	20/55 (36.4)	18/57 (31.6)

Based on Table 1, Appendix E, CS[23]
Percentages are based on the treatment group N (overall study population, age groups, or BMI groups)
BMI = body mass index; CS = company submission; DB = double-blind; n = number of subjects with at least 1 event; TEAE = treatment-emergent adverse event

3.2.6.5.2 Study 303[18]

The effect of daridorexant on TEAEs was found to be consistent across all subgroups (Table 3.49).

Table 3.49: Subjects with at least one TEAE during the double-blind study period by subgroup for age, BMI, sex and race

	Daridorexant 50 mg; n/N (%)	Placebo; n/N (%)
Overall study population	[redacted]	[redacted]
Age at screening (years)		
< 65	[redacted]	[redacted]
≥ 65	[redacted]	[redacted]
Age at screening (years)		
< 75	[redacted]	[redacted]
≥ 75	[redacted]	[redacted]
BMI at screening (kg/m^2)		
< 25	[redacted]	[redacted]
25–30	[redacted]	[redacted]
> 30	[redacted]	[redacted]
Sex		
Male	[redacted]	[redacted]
Female	[redacted]	[redacted]
Race		
White	[redacted]	[redacted]
Black or African	[redacted]	[redacted]
Other	[redacted]	[redacted]

Based on Table 1, Appendix E, CS[23]
Age group and BMI group were determined at screening of the confirmatory Study (301 or 302).

	Daridorexant 50 mg; n/N (%)	Placebo; n/N (%)
BMI = body mass index; CS = company submission; TEAE = = treatment-emergent adverse event		

3.2.7 Included studies: Supporting evidence

Not applicable.

3.2.8 Ongoing studies

Not applicable.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

No indirect comparison and/or multiple treatment comparison was carried out.

EAG comment:

- An indirect treatment comparison (ITC) is required for this submission, to allow an estimation of daridorexant versus ECM (the decision problem comparison) from the current data on daridorexant versus placebo and other data on ECM versus placebo. Without such an analysis it is difficult to see how the submission can adequately address the decision problem.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

Not applicable.

3.5 Additional work on clinical effectiveness undertaken by the EAG

The EAG carried out its own independent risk of bias assessments for studies 301[15] and Study 303.[16], using the information from the respective CSR documents. These confirmed that the risk of bias in both studies was low (see section 3.2.4).

The EAG also carried out between-arm statistical analyses where these had not been performed by the company. These were largely for results pertaining to study 303[16]. For clarity, these results have been integrated with the company's results, but have been clearly signposted, as well as being differentiated by being placed within the EAG comments.

3.6 Conclusions of the clinical effectiveness section

The CS[1] and response to clarification[3] provided sufficient details for the EAG to appraise the literature searches conducted to identify studies of the efficacy and safety of pharmacological treatments for insomnia disorder. Searches were conducted in March 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases and trials registers were searched. Overall, the EAG has no major concerns about the literature searches conducted, however separate adverse events searches and a broader approach to conference searching may have retrieved additional studies.

The NICE scope and decision problem involved evaluation of daridorexant against established clinical practice. However, the company only provided evidence for daridorexant against placebo, without any attempt to compare daridorexant to established practice using indirect treatment comparisons. It is therefore difficult to evaluate daridorexant in the appropriate context of the decision problem.

The evidence submitted lacked a key paper and was therefore incomplete. The two included studies were high quality RCTs, and internal validity was judged to be good. These studies suggest that daridorexant yields clinical benefits compared to placebo in the short term (3 months) but that in the longer term these benefits may become less certain. For example, whilst sWASO, sTST and sLSO were all significantly improved by daridorexant compared to placebo at 3 months, this was no longer the case at 12 months. It should be noted that the higher levels of uncertainty were only apparent after the EAG had carried out their own between-arm analyses, which had not been carried out by the company for all analyses. The EAG accepts that the uncertainty may be partly due to the lower statistical power of the longer-term study, but it cannot be assumed that this is the sole cause.

Adverse events appeared to be generally non-serious, and therefore less likely to have a significant negative impact on any benefits of daridorexant. It is notable that daridorexant had a significantly lower risk of rebound insomnia in terms of WASO and LPS than placebo.

In terms of applicability, questions remain about the relevance of the study findings to the UK population. Although uncertain, there was a possible difference in the proportions of ethnicity groups between the target UK population and the two included studies. There is additional uncertainty about whether ethnicity is an important factor influencing outcomes: Study 301 did not sub-group for ethnicity, and whilst Study 303 did not find evidence that ethnicity was an effect modifier, analyses were only presented for two outcomes. Although there is no clear evidence that ethnicity *is* an effect modifier, there is insufficient evidence to support the company's claim that ethnicity is *not* an effect modifier. In addition, the study populations were largely naïve to the main alternative treatment CBT-I. This creates a serious divergence from the intended clinical population for daridorexant: people who have not responded to CBT-I.

4. COST EFFECTIVENESS

4.1 EAG comment on company's review of cost effectiveness evidence

Three SLRs were performed with the objectives to identify and select relevant 1) cost effectiveness analysis (CEA) studies (CS Appendix G[24]); 2) HRQoL studies (CS Appendix H[25]); 3) costs and healthcare resource use studies (CS Appendix I[26]).

4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.[1] The CADTH evidence-based checklist for the PRESS, was used to inform this critique.[8, 9] The CS[1] was checked against the STA specification for company/sponsor submission of evidence.[10] The EAG has presented only the major limitations of each search strategy in the report.

Appendix G of the CS provides details of a SLR conducted to identify economic evaluations of therapies for chronic insomnia disorder.[24]

A summary of the sources searched is provided in Table 4.1.

Table 4.1: Data sources searched for economic evaluations (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE and MEDLINE In-Process	Ovid	All*	6/4/22
Embase	Ovid	All*	6/4/22
CENTRAL	Ovid	All*	6/4/22
CDSR	Ovid	All*	6/4/22
NHS EED	Ovid	All*	6/4/22
EconLit	Ovid	All*	6/4/22
PsycINFO	Ovid	All*	6/4/22
Conferences			
European Sleep Research Society	Internet	2020-2022	Not stated
Sleep Europe			
ISPOR			
ISPOR Europe			

*The CS and response to clarification state that no date limit was applied, however it is not clear which database segment was used as the database start and end dates were not supplied[1, 3]
CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; EED = Economic Evaluations Database; ISPOR = International Society for Pharmacoeconomic and Outcomes Research; NHS = National Health Service; SFRMS = Société Française de Recherche et Médecine du Sommeil

Appendix H of the CS provides details of a SLR conducted to identify the humanistic burden of chronic insomnia disorder.[25]

A summary of the sources searched is provided in Table 4.2.

Table 4.2: Data sources searched for HRQoL studies (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE and MEDLINE In-Process	Ovid	All*	6/4/22
Embase	Ovid	All*	6/4/22
CENTRAL	Ovid	All*	6/4/22
CDSR	Ovid	All*	6/4/22
NHS EED	Ovid	All*	6/4/22
EconLit	Ovid	All*	6/4/22
PsycINFO	Ovid	All*	6/4/22
Conferences			
European Sleep Research Society	Internet	2020-2022	Not stated
Sleep Europe			
ISPOR			

*The CS and Response to Clarification state that no date limit was applied, however it is not clear which database segment was used as the database start and end dates were not supplied[1, 3]
CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; EED = Economic Evaluations Database; ISPOR = International Society for Pharmacoeconomic and Outcomes Research; NHS = National Health Service; SFRMS = Société Française de Recherche et Médecine du Sommeil

Appendix I of the CS provides details of a SLR conducted to identify the relevant studies evaluating the costs and healthcare resource utilisation (HCRU) for patients with chronic insomnia disorder.[26]

A summary of the sources searched is provided in Table 4.3.

Table 4.3: Data sources searched for cost/resource use studies (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE and MEDLINE In-Process	Ovid	All*	6/4/22
Embase	Ovid	All*	6/4/22
CENTRAL	Ovid	All*	6/4/22
CDSR	Ovid	All*	6/4/22
NHS EED	Ovid	All*	6/4/22
EconLit	Ovid	All*	6/4/22
PsycINFO	Ovid	All*	6/4/22
Conferences			
European Sleep Research Society	Internet	2020-2022	Not stated
Sleep Europe			
ISPOR			
ISPOR Europe			

Resource	Host/Source	Date Ranges	Date searched
*The CS and Response to Clarification state that no date limit was applied, however it is not clear which database segment was used as the database start and end dates were not supplied[1, 3] CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; EED = Economic Evaluations Database; ISPOR = International Society for Pharmacoeconomic and Outcomes Research; NHS = National Health Service; SFRMS = Société Française de Recherche et Médecine du Sommeil			

EAG comment:

- Searches were undertaken in April 2022 to identify economic, HRQoL and healthcare resource use/cost data on chronic insomnia disorder. The CS, Appendices G, H and I and the company's response to clarification provided sufficient details for the EAG to appraise the literature searches.[1, 3, 24-26]
- A good range of databases was searched.
- Database search strategies contained a population facet for insomnia and sleep disorders and the relevant measurement tools. This facet was then combined with terms for daridorexant and its comparators. This facet also included additional search terms for cognitive behavioural therapy that were not in the clinical effectiveness searches.
- Searches were well structured, transparent and reproducible, and a good range of subject indexing terms (MeSH/EMTREE) and free text was used.
- Overall, a good range of subject indexing terms (MeSH/EMTREE) and free text was used.
- The database searches for economic evaluations had a 2002 publication date limit, and the HRQoL and resource use/cost searches had a 2012 publication date limit.
- Conference proceedings searches were conducted for the two most recent meetings available for four named conferences. However, conference proceedings were excluded from the Embase search, which can be a useful source of additional conference papers. In response to clarification, the Company stated that: 'Conference proceedings were excluded from the Embase clinical effectiveness searches due to a high volume of yield resulting from any conference proceedings reporting on 'insomnia', introducing a high number of irrelevant publications to screen. Hence, a targeted approach was followed by specifically hand searching conferences of interest in the past two years. It is standard practice to search for conference proceedings of preceding two years, as any study results published before would be reported in a peer review publication, which can be captured through database search.' However, the exclusion of conference proceedings from the Embase searches only removed around 240 references from the economic evaluations search, around 230 references from the HRQoL search and around 850 references from the resource use/cost search. After deduplication, this would not have been likely to have greatly increased the screening burden. Amongst these results were references from the World Sleep Congress and the Annual Meeting of the Associated Professional Sleep Societies and more generic neurology conferences, which may have provided additional useful references. The EAG notes also that it is not necessarily the case that all conference proceedings will be published in peer reviewed journals.
- Database searches for economic evaluations were limited to English language publications only. The HRQoL and resource use/cost searches had no language limit.
- Study design filters were used in the databases searches to identify relevant economic evaluations, HRQoL/utilities studies and healthcare resource use/cost data. The study design filters were not referenced, so it was unclear whether the filters used were published objectively-derived filters. The filters contained a combination of subject heading terms and free text terms and the EAG deemed them to be adequate. However, the EAG notes the use of filters in the NHS EED and EconLit

searches. As these databases contain only economic and related publications, the EAG believes it was not necessary to include filters in these strategies, and this may have resulted in unnecessarily restricting the results retrieved. Given the low numbers of records found before the filters were applied it may therefore have been advisable to search these databases without study design filters.

4.1.2 Inclusion/exclusion criteria

Inclusion and exclusion criteria for the review on cost effectiveness studies, HRQoL studies and costs and resource use studies are presented in Table 4.4.

Table 4.4: Eligibility criteria for the SLRs

	Inclusion criteria	Exclusion criteria
Patient population	Adults ≥18 years old with chronic insomnia disorder	Paediatric (<18 years old) patients Patients with acute (short-term) insomnia
Intervention	Individual pharmacological interventions (see Table 8 in Appendices G, H and I) CBT-I Combination therapy	Non-pharmacological interventions apart from CBT-I Barbiturates, chloral hydrate, ethchlorvynol and quetiapine Herbal products and medical devices
Comparator	Any or none	
Outcomes(s) 1 (Published economic evaluations)	Cost effectiveness/utility analysis results such as ICER and QALYs Cost-minimisation analysis results Cost-benefit analysis results Budget impact model results	Publications that do not report data on relevant outcomes
Outcomes(s) 2 (HRQoL studies)	HRQoL (e.g., EQ-5D, SF-36) Utilities/disutilities	Publications that do not report data on relevant outcomes
Outcomes(s) 3 (Cost/resource use studies)	Direct or indirect costs of treatment or illness Resource use (hospitalisations, length of stays, ER visits) Drivers of cost/resource use (healthcare, hospital, drug related)	Publications that do not report data on relevant outcomes
Study design 1 (Cost effectiveness analysis studies)	Cost effectiveness analyses, cost-utility analyses, cost-minimisation analyses, and cost-benefit analyses Budget impact models	Clinical trials, observational/real-world studies, case reports, non-systematic reviews, commentary, and letters
Study design 2 (HRQoL studies)	Observational/real-world studies (prospective and retrospective) HRQoL studies	Cost effectiveness analyses, cost-utility analyses, cost-minimisation analyses and cost-benefit analyses Budget impact models

	Inclusion criteria	Exclusion criteria
	Utility studies RCTs only for HRQoL data	Case reports, non-systematic reviews, commentary and letters
Study design 3 (Cost/resource use studies)	Observational/real-world studies (prospective and retrospective) Cost-of-illness studies	Cost effectiveness analyses, cost-utility analyses, cost-minimisation analyses, and cost-benefit analyses Budget impact models Clinical trials Case reports, non-systematic reviews, commentary, and letters
Based on Table 8, Appendix G,[24] Table 8, Appendix H,[25] Table 8, Appendix I[26] CBT-I = Insomnia-related cognitive behavioural therapy; EQ-5D = European Quality of Life-5 Dimensions; ER = emergency room; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; RCTs = randomised controlled trials; SF-36 = 36-item short form; SLRs = systematic literature reviews		

EAG comment:

- The EAG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies.

4.1.3 Conclusions of the cost effectiveness review

The CS provides an overview of the included cost effectiveness, utility and resource use and costs studies, but no specific conclusion was formulated.

The CS[1] and response to clarification[3] provided sufficient details for the EAG to appraise the literature searches conducted to identify economic, HRQoL and cost data on chronic insomnia disorder. Searches were conducted in April 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases were searched. Overall, the EAG has no major concerns about the literature searches conducted, although a broader approach to conference searching may have retrieved additional studies.

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 4.5: NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Consistent with reference case
Perspective on costs	NHS and PSS	Consistent with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	Consistent with reference case
Time horizon	Long enough to reflect all important differences in costs	Unclear whether all relevant costs and effects are captured

Element of health technology assessment	Reference case	EAG comment on company submission
	or outcomes between the technologies being compared	within the 12-month time horizon
Synthesis of evidence on health effects	Based on systematic review	Partly consistent with reference case (no review is used to identify relevant mapping functions or sources that could potentially be used to develop a mapping function)
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Consistent with reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Consistent with reference case
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Unclear whether the UK tariff was used in the NWHS population (used for developing the mapping algorithm)
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Consistent with reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Consistent with reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Not applicable (given the 12-month time horizon)

EAG = Evidence Assessment Group; EQ-5D = European Quality of Life-5 Dimensions; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NWHS = National Health and Wellness Survey; PSS = Personal Social Services; QALYs = quality-adjusted life years; UK = United Kingdom

4.2.2 Model structure

The economic model (Microsoft Excel) estimated effectiveness through mapping the observed ISI© scores to EQ-5D utility values. For this purpose, the company developed a mapping algorithm. More details regarding the estimation of utility values are provided in Section 4.2.8 of this report. Costs were estimated for the following categories: treatment costs and medical costs (GP, emergency room attendances, inpatient care). More details regarding the estimation of resource use and costs are provided in Section 4.2.9 of this report.

No impact of AE on estimated effectiveness and costs was assumed.

The economic model consisted of multiple regression models to estimate costs and effects for months 0, 1, 3,6, 9 and 12 (based on observed ISI© scores from Study 301 and Study 303). In the CS base case, estimated costs and effects were restricted to the observed data period (i.e., no extrapolation is applied).

EAG comment:

- The main concerns of the EAG relate to the definition of the model type. The model type (e.g., decision tree, state-transition model) was not clearly specified in the CS. In response to clarification question B1, the company described it as a 'mediated' analysis. Moreover, the company indicated "We are not aware that any formal terminology has entered the lexicon, which was why we did not state the model form." In response to clarification question B2a regarding model-based and trial-based approaches, the company recognises the "that the model is something of a hybrid" … "perhaps a hybrid trial-model". Although not common, the EAG believes the company's approach is reasonable.

4.2.3 Population

The population as defined by the NICE scope and as described in the CS (Section B.1.1, Table 1) is "Adults with insomnia disorder", without further specification in the cost effectiveness section of the CS. Upon request for clarification, the company confirmed that the modelled population was identical to the population enrolled into Study 301 (i.e., adults with insomnia disorder as per the DSM-5® criteria and with ISI© score ≥15).

Tables 7 and 27 of the CS describe the in- and exclusion criteria used for Study 301 and Study 303. The exclusion criteria for both studies contain criteria regarding the presence or history of mental health diseases.

EAG comment:

- There are two discrepancies between the scope listed in Study 301 and Study 303, and UK clinical practice leading to uncertainty around the generalisability of the model. Firstly, Study 301 and Study 303 excluded patients with conditions related to mental health problems (acute or unstable psychiatric conditions). As insomnia and other mental health disorders are frequently comorbid[27], this may have excluded a considerable part of the treatment population as it would be seen in practice. The company argued that the study results could be extrapolated to patients with a comorbid mental health problem because the underlying mechanism for insomnia is the same. Evidence to justify this claim was not provided. Further, if the underlying mechanism is the same, this does not mean that the effectiveness between the populations must be equal. Moreover, ingesting antidepressant medication such as selective serotonin reuptake inhibitors (SSRIs) frequently leads to sleep problems as a side-effect[28]. Treatment for mental health disorders may therefore be a confounder of unknown size. Secondly, CBT-I was allowed in the treatment populations of both Study 301 and Study 303. However, the company insisted that daridorexant would be given only as an alternative to CBT-I, if CBT-I failed, was inaccessible or was refused by the patient. The EAG concludes that the exclusion of patients with mental health problems and the inclusion of patients receiving CBT-I in Study 301 and Study 303 results in uncertainty surrounding the generalisability of the treatment effect to the anticipated treatment population.

4.2.4 Interventions and comparators

The intervention considered in the cost effectiveness analysis was daridorexant 50 mg. The company clarified that the 25 mg and 50 mg dosages were included in the anticipated market authorisation.

According to the CS, treatment duration should be as short as possible, with the appropriateness of continuing treatment being assessed within 3 months and periodically thereafter. However, no stopping rule is explicitly mentioned within the CS.

No-treatment was used as the comparator in the cost effectiveness analysis, and the placebo arm of the trial serves as a proxy for no-treatment based on the analysis of Study 301[1]. In contrast, the NICE scope (CS Table 1) specified that "established clinical management (including sleep hygiene advice) without daridorexant" would be the most appropriate comparator[1]. Multiple comparators are mentioned in the NICE clinical knowledge summary[29] including sleep hygiene advice, CBT-I, non-benzodiazepine hypnotic medication, zolpidem, zopiclone, benzodiazepines and melatonin. In document A of the CS[30], the company shows multiple locations in the treatment pathway in which the use of daridorexant may be used and as a result which comparators are appropriate. The company justified the selection of no-treatment as a comparator by stating the following: "Daridorexant is the first insomnia treatment with longer term data for the treatment of insomnia disorder". In the clarification response, the company added that "none of the currently approved pharmacological treatments are recommended for long-term use"[3].

EAG comment:

- The main concerns of the EAG relate to: a) the use of no-treatment as comparator, and b) the lack of evidence provided for the use of the 25 mg dosage.

    a) The NICE scope specified "established clinical management (including sleep hygiene advice)" as the comparator in this submission. In addition, the NICE CKS mentions sleep hygiene advice, CBT-I, non-benzodiazepine hypnotic medication, zolpidem, zopiclone, benzodiazepines and melatonin. In contrast, the company applied no-treatment as the comparator in their health economic model, justified by stating that daridorexant is the first insomnia treatment with longer term data available and the only pharmacological treatment recommended for longer term use. Though not clearly defined in the CS, "long term" seems to refer to a duration of approximately >4 weeks (CS Section B.1.3.5). The company did not comply with the EAG's request to add other comparators listed in the NICE scope. Although the EAG acknowledges that other potential comparators may not be approved for long term use, the EAG finds that the comparison with short-term use of a drug is nonetheless relevant. In response to clarification question A8, the company responded that CBT-I was not included because daridorexant should be given when CBT-I was inaccessible, patients were unable to follow CBT-I or CBT-I was refused by the patient. Further, Figure 1 in Document A[30] of the CS describes daridorexant as an alternative first-line treatment if CBT-I is inaccessible, patients are unable to follow or refuse CBT-I. As daridorexant is described as an alternative treatment, the EAG believes that CBT-I should be included as a comparator to daridorexant. Further, the scope clearly states that, at least sleep hygiene advice should have been included as a comparator. The EAG concludes that without the inclusion of all relevant comparators, the model lacks relevance to the decision-problem.

    b) Although the anticipated market authorization for daridorexant includes the 25 mg dosage, the company did not include the 25 mg dosage in the cost effectiveness model. The company claimed that the 25 mg dosage would likely only be used where there is co-administration of moderate CYP3A4 inhibitors. However, in both Study 301 and Study 303 individuals taking CYP3A4 inhibitors were excluded from participating. Hence, the question remains whether the

results of Study 301 or Study 303 can be used to inform the use of 25 mg dosage in this specific population.

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and PSS perspective with a 12-month time horizon. Discount rates were not applicable given the 12-month time horizon. Discount rates of 3.5% for both effects and costs were used in the lifetime model scenario.

EAG comment:

- The main concerns of the EAG relate to the model time horizon of 12 months. The company indicated (response to clarification question B3a[3]) that [redacted] remain on treatment at the end of the 12-month time horizon. Therefore, it is questionable whether all relevant costs and benefits are captured within this period.

4.2.6 Treatment effectiveness and extrapolation

The main sources of evidence on treatment effectiveness used for the intervention and comparator were the ISI© scores, an exploratory trial outcome from the daridorexant 50 mg arms of Study 301 (N=310; NCT03545191)[15] and Study 303 (N=137) (NCT03679884)[18]. Both studies were multi-centre, randomised, double-blind, placebo-controlled, parallel-group phase III studies that evaluated daridorexant in subjects with insomnia disorder for 12 weeks (Study 301) and 40 weeks (Study 303). Study 303 was an extension of confirmatory studies 301[15] and 302[16].

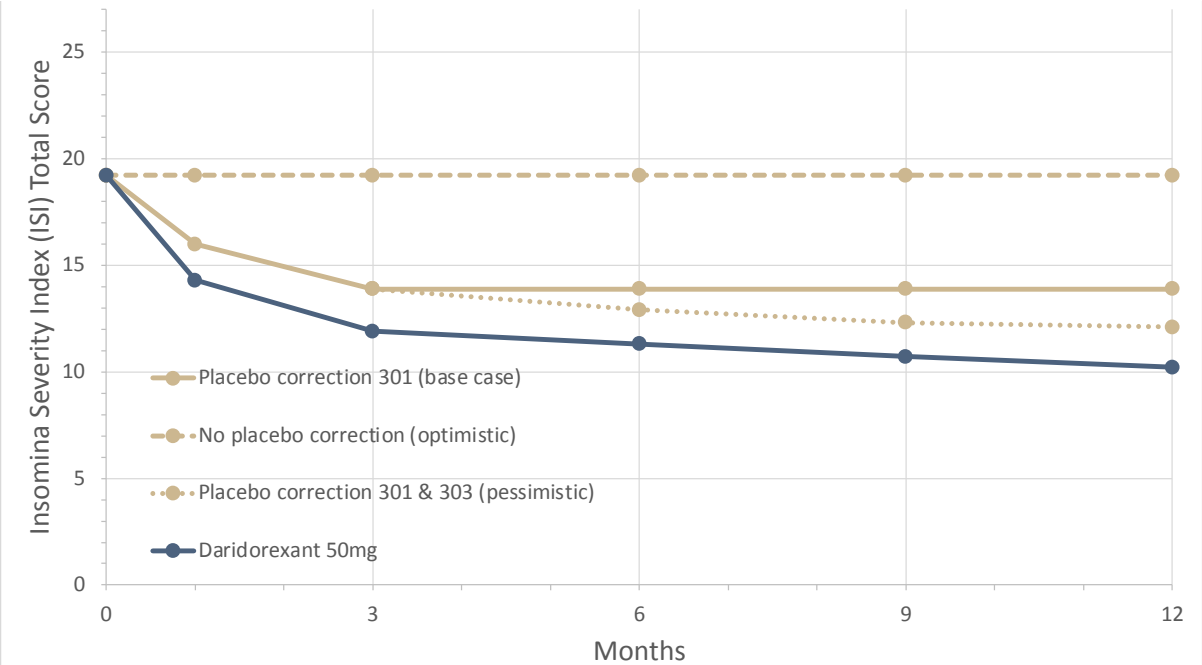
The effectiveness of daridorexant for the first 12 weeks was estimated via seemingly unrelated regression (SUR) of ISI© scores derived from Study 301 to adjust for both baseline ISI© and placebo effects. Although not clearly described in the CS, from the model file it appears that for the 6th, 9th and 12th month, the effectiveness of daridorexant was modelled by adding the treatment effect observed in Study 303 in each respective month (i.e., 6th, 9th, and 12th month) to the adjusted value of the 3rd month derived from the SUR calculations. For the no-treatment arm in the company's base case, placebo was adjusted based on Study 301, as the company argued that Study 303 presented evidence of selective attrition in both arms[18].

Treatment discontinuation was based on the observed discontinuation rates from both studies but, according to the company, only incorporated for the daridorexant arm. The company justified this by assuming that no-treatment patients could not dropout from receiving no treatment. Additionally, the company included two extra assumptions to model the impact of discontinuation: (1) discontinuation occurred at the midpoint of the studied periods; (2) treatment costs were incurred for the full period assuming that prescriptions would be filled at the start of the period before discontinuation occurs.

In addition to the base case analysis, the company performed optimistic and pessimistic scenario analyses (as labelled by the company). In Figure 4.1 below, the blue line titled "Daridorexant 50 mg" represents the modelled ISI© scores for the treatment arm, based on the analysis of Study 301 and Study 303. For the base case, it was assumed that patients receiving the intervention would experience the adjusted improvement observed in Study 301 and Study 303 (see above) and that patients receiving no-treatment would not experience the improvement after 3 months as observed in the placebo arm in Study 303 (i.e., would continue with the same ISI© score from the end of Study 301). For the optimistic scenario, the company assumed that patients receiving no-treatment would keep the ISI© score of placebo arm at baseline for the complete time horizon (i.e., not improving their ISI© score at all for the

12 months). For the pessimistic scenario, patients receiving no-treatment were modelled to experience the improvement in the ISI© scores that were observed for placebo patients in studies 301 and 303.

Figure 4.1: Modelled trajectory of ISI© in different scenarios (Based on CS, Figure 15)



CS = company submission; ISI© = Insomnia Severity Index

EAG comment:

- The main concerns of the EAG relate to: a) justification for the use of the ISI© score for the estimation of treatment effectiveness; b) application of the placebo adjustment, c) use of seemingly unrelated regression; and d) assumptions for treatment discontinuation.

  a) Treatment effectiveness in the economic analyses was based on ISI© scores, an exploratory trial outcome, from studies 301 and 303. Other clinical primary (i.e., WASO and LPS), and secondary (i.e., sTST and IDSIQ) efficacy endpoints were collected in the trials but were not used to inform treatment effectiveness. The company justified the use of ISI© scores for the estimation of treatment effectiveness due to the complexity of assessing treatment outcomes in insomnia disorder and the lack of mapping algorithms to the EQ-5D for other outcomes. Nonetheless, as per CS, there were no mapping algorithms for deriving EQ-5D utilities from ISI© scores, and, hence, the company developed a novel mapping algorithm. The company also stated that ISI© scores should be prioritised given that it was "the key effectiveness parameter of the economic model". However, the EAG considers that this line of argumentation is flawed, as the choice to use an outcome in the economic model does not make it automatically the best outcome to model relative effectiveness. In addition, the company justified this choice based on the "equivalent" results of the number needed to treat (NNT) analysis, comparing different outcomes of Study 301 (Table 5, clarification response)[3]. However, these analyses were based only on the results of Study 301 (with three months of follow up). The follow- up in Study 303 (40 weeks) was longer than in Study 301 (12 weeks). Hence, the use of data only from Study 301 may be unrepresentative of the overall effect of the intervention. Moreover, the EAG is not familiar with the assumption of comparability between endpoints based on similarity of NNTs, as it is not a common practice. The same NNT does not automatically confirm comparability,

true comparability between endpoints would be manifested not just by the same NNT but also by correlation between events across endpoints. Given the before-mentioned arguments, the EAG considers that the use of ISI© scores was not sufficiently justified.

b) The company included the placebo effect reported both in Study 301 and Study 303 only in their pessimistic scenario analysis (this increased the ICER to £[redacted] and £[redacted] including the dropout rates). In the company's base case, placebo was only adjusted for the first 3 months in the no-treatment arm, but not thereafter. The EAG had two main issues with the placebo adjustment assumptions and application:

   i. The company stated that the base case should not include the placebo effect observed in Study 303 (study that based the results from the 3rd month to the 12th month), arguing that Study 303 "presented evidence of selective attrition". However, as per clarification question B11 and CS, Figure 13b, selection attrition was also present in the treatment arm (and even had a larger effect than in the placebo arm).

   ii. For the EAG, it was unclear whether improvements in ISI© scores would be attributable to 'natural improvement' of the symptoms, regression to the mean, or placebo effect in the placebo group. When asked to clarify about the possibility of regression to the mean, the company argued that regression to the mean could not be responsible for the observed effect in the placebo because: (1) There was an initial screening phase in the design of the study, followed by randomisation at visit 4 (20-31 days later). Nonetheless, at clarification question B9a, the company agreed that the lower ISI© values at randomisation could be attributable to regression to the mean[3]. (2) There was a rebound effect between studies 301 and 303. Despite the fact that there was a rebound effect between the end of Study 301 (12th week) and the beginning of Study 303 (16th week), Study 303 continued for 40 weeks more, and in those weeks patients could have improved their scores naturally, especially given than insomnia is highly related to lifestyle factors, as mentioned in CS B.1.3.2[1].

Therefore, the EAG considers that (1) the effect of selective attrition on the daridorexant group and (2) the possibility of regression to the mean, were not sufficiently justified by the company, and these effects could have biased the comparison in favour of the intervention. Hence, the EAG will incorporate placebo adjustment for the 12 months in the EAG base case. Moreover, as discussed in Section 4.2.4., the use of no-treatment (i.e., no relevant comparators) may lack relevance for the decision-problem and may not appropriately represent UK clinical practice all over.

c) The company used the SUR procedure to model the relationship between ISI© scores at month 1 and month 3 (i.e., Study 301); however, the company did not perform the same analysis for Study 303, as their modelling team did not have access to the patient-level data of said study[3]. However, the EAG notes that the company is the sponsor of that study. Two modelling approaches were used to obtain the ISI© scores used in the model, one for the data of Study 301 (i.e., SUR), and one for the data of Study 303. In the base case, ISI© scores from the 3rd month onwards (i.e., 6th, 9th, and 12th month) were calculated based on the ISI© score from the 3rd month plus the treatment effect of daridorexant 50 mg from Study 303 at each time-point, rather than using SUR. The EAG considers that the use of two different methodologies was insufficiently justified by the company and the EAG cannot fully assess the potential uncertainty introduced by these different methodological choices. Likewise, the EAG would like to highlight that the data limitations do not provide justification for the methodological choices used.

d) Treatment dropout was based on the observed dropout rates from studies 301 and 303, but, as per CS, treatment dropout was only incorporated in the intervention arm. When asked to provide how many patients were modelled to discontinue treatment, the company provided the number of patients dropping out at each time point in both studies in response to clarification (Table 11, clarification response). The EAG found multiple issues with the application of dropout rates:

   i. According to the company, treatment discontinuation (i.e., dropout rates) were not incorporated in the no-treatment group as patients receiving no treatment would not be able of dropout from the lack of treatment. However, in the economic model provided by the company, dropout rates are applied to the incremental values (i.e., the difference between the daridorexant and no-treatment group), instead of being applied to the daridorexant group alone. The EAG will explore applying only the dropout rates to the daridorexant group in their base case, as defended in the CS. [1]

   ii. As the company considered dropout rates from Study 303 were [redacted] clinical practice more accurately than the ones from Study 301, the company provided a scenario analysis using "similar" dropout rates in the first and third month (the ones which were based on Study 301). For the scenario analysis, the company used [redacted] dropout rate for the 1st month (instead of 4%), [redacted]% for the 3rd month (instead of 5%) and [redacted]% for the 6th , 9th, and 12th months, instead of the dropout rates from Study 303 (i.e., [redacted]%, [redacted]%, and [redacted]%, respectively). The EAG considered that the assumption of the [redacted]% and [redacted]% discontinuation rates, was not sufficiently justified, as the calculation was based on that [redacted]% is [redacted]% (i.e., the dropout rate at 14 weeks); likewise, the company did not provide sufficient justification on using a lower dropout rate for the 6th and 12th months, when those were supposed to "reflect clinical practice more accurately". Hence, the EAG will explore in a scenario analysis. the use of similar dropout rates of Study 303 in Study 301 (i.e., [redacted]% and [redacted]%) but keeping the same dropout rates from Study 303 for the 6th, 9th and 12th month (i.e., [redacted]%, [redacted]%, and [redacted]%, respectively).

4.2.7 Adverse events

Adverse events were not included in the model. According to the company, the difference in AEs rates between the two arms (which were reported in ≥2% of the treatment groups) were of a mild nature and indicated a favourable safety profile for daridorexant 50 mg as seen in Table 4.6. Therefore, the company assumed a negligible effect of AEs on HRQoL and costs.

Table 4.6: AEs reported at least once in either treatment group

Treatment-emergent adverse event	Study 301		Study 303	
	Daridorexant 50 mg N=308; n (%)	Placebo N=309; n (%)	Daridorexant 50 mg N=137; n (%)	Placebo N=128; n (%)
Subjects with at least one TEAE	116 (37.7)	105 (34.0)	[redacted]	[redacted]
Subjects with at least one SAE	3 (1.0)	7 (2.3)	[redacted]	[redacted]
Subjects with at least one AESIs	2 (0.6)	1 (0.3)	[redacted]	[redacted]

Treatment-emergent adverse event	Study 301		Study 303	
	Daridorexant 50 mg N=308; n (%)	Placebo N=309; n (%)	Daridorexant 50 mg N=137; n (%)	Placebo N=128; n (%)
Based on CS Table 18, 19, 20, 38, 39, & 40 [1] AESI = adverse events of special interest; CS = company submission; TEAE = treatment emergent adverse event, SAE = serious adverse event				

EAG comment:

- The main concern of the EAG relates to the exclusion of AEs from the cost effectiveness model.
- According to the CS, during the double-blind study period in Study 301 and Study 303, 37.7% and 38.0% of the subjects reported at least one TEAE in the daridorexant 50 mg arm respectively, while 1.0% and 5.1% of the participants experienced at least one treatment-emergent SAE (Table 4.6). In their response to clarification question B12, the company argues that the side effects are minor and not that different from placebo, therefore, it would not be expected to have consequences on resource use or HRQoL. Moreover, the company provided a scenario in which they assumed a mild nature AE with 0.2 utility decrements and a cost of £5, which lead to an increase in the base case ICER from [redacted] to [redacted]. The EAG appreciates the scenario analysis conducted by the company; however, it considers the assumptions underlying it to be simplistic. Hence, although the EAG does not expect a large impact on the cost effectiveness results, it prefers, as requested in the clarification letter, an updated cost effectiveness model and scenario analyses incorporating all AEs from Study 301 and Study 303 as well as the impact on estimated costs and effects.

4.2.8 Health-related quality of life

The company stated that EQ-5D was not collected in clinical studies and the SLR did not identify any HRQoL studies relevant to the cost effectiveness model. Therefore, a novel mapping algorithm based on the NHWS dataset was used to map ISI© data from Study 301 and Study 303 to EQ-5D values. A generalised linear model with a gamma distribution family and log link function, including one - utility as the dependent covariate, was used to create the mapping function. The fitted mapping function based on NHWS data resulted in the following model to estimated utility values: $E[U] = 1 - \exp\{-1.849 + 0.047 \times ISI©\}$. The CS did not include a detailed description and justifications with regards to the mapping of ISI© scores to derive EQ-5D utilities, e.g., it was unclear whether the clinical and demographic characteristics of people in the estimation sample (NHWS) were similar to the characteristics of the target sample (studies 301 and 303). In addition, the CS did not include details regarding the conceptual overlap between the ISI© and EQ-5D dimensions/instruments and it was not reported whether the mapping function was validated and whether other model types were considered (more suitable) for estimating the mapping algorithm.

The resulting utility values for ISI© scores at different timepoint are summarised in Table 4.7.

Table 4.7: Health state ISI© scores and utility values

Time (months)	No-treatment		Daridorexant	
	ISI©	Utility	ISI©	Utility
0	19.212	0.613	19.212	0.613
1	15.986	0.667	14.290	0.693
3	13.885	0.698	11.903	0.725
6	13.885	0.698	11.300	0.733

Time (months)	No-treatment		Daridorexant	
	ISI©	Utility	ISI©	Utility
9	13.885	0.698	10.700	0.740
12	13.885	0.698	10.200	0.746
Based on Economic model ISI© = Insomnia Severity Index				

4.2.8.1 Disutility values

No (AE) disutilities were applied to the economic model.

EAG comment:

- The main concerns of the EAG relate to a) several issues regarding the mapping of ISI© scores reported in Study 301 and Study 303 to derive EQ-5D utilities, and b) model type for estimating the mapping algorithm.

  a) The company used a generalised linear model to create a mapping function from the cross-sectional NHWS survey to derive EQ-5D utilities from ISI© scores reported in studies 301 and 303. The EAG is concerned about the following issues:

Firstly, NICE DSU TSD 10[31], guidance regarding the use of mapping methods to estimate health state utility values (HSUVs), states that "*In order to be confident about the generalisability of the mapping function to the target sample, the clinical and demographic characteristics of people in the estimation sample should be as similar to the characteristics of the 'target' sample to which the mapping algorithm will be applied as possible*". However, as confirmed by the company in response to clarification, there was few overlaps between characteristics of patients in the NHWS dataset and patients in Study 301 and Study 303. One important difference highlighted by the company was the difference in mean ISI© score, which was 12.6 (subclinical insomnia) in the NHWS population and 19.2 (moderate insomnia) in Study 301. The company argued this to be a positive attribute since a broader range of ISI© and EQ-5D values should results in a more robust mapping algorithm. The EAG, however, questions the generalisability of the mapping function to the target population (i.e., Study 301 and Study 303) and hence considers the ability of the mapping algorithm to predict the utility values of the patients in the Study 301 and Study 303 to be uncertain.

Secondly, the EAG is concerned about the (lack of) conceptual overlap between the ISI© and EQ-5D instruments. A recent review[32], of mapping studies found that explanatory power using R-squared was often low for models that involved mapping a condition-specific measure onto a generic preference-based measure (e.g., ISI© to EQ-5D). This may occur due to limited conceptual overlap as important dimensions in the ISI© instrument may not appear in the EQ-5D instrument and vice versa. In response to the clarification letter, the company states that the ISI© correlates with the EQ-5D and is suitable to estimate the QALYs presented in the submission. However, the company acknowledges that it is very plausible that the EQ-5D does not fully capture the impact of insomnia disorder on HRQoL and considers it reasonable that the benefits of daridorexant on HRQoL is currently underestimated. The EAG further questioned why other clinical outcomes (i.e., WASO, LPS, sTST and IDSIQ scores) reported in the pivotal trials were not utilised to map to the EQ-5D, to which the company replied that there were no available data sources to estimate a mapping function for these outcomes. Although the EAG acknowledges this data availability limitation, concerns regarding the conceptual overlap between the ISI© and EQ-5D instruments and their suitability to estimate HRQoL in insomnia remains. A correlation matrix analysis to assess the correlation within dimensions of both instruments could be

provided to consider convergent validity (the degree to which a dimension correlates with another dimension measuring the same concept).

Thirdly, NICE DSU TSD 10 states that a validation stage should also be applied, whereby the regression results are validated against another dataset. However, in response to clarification, the company argued that it was not clear that any external data were available and hence the regression results were not validated. As the EAG considers the validation step to be important, it alternatively suggests the company to validate their regression results by randomly splitting their estimation dataset into an estimation sample and a validation sample. The mapping function can then be estimated on the estimation sample and its performance can be examined using the validation sample (more details are reported in NICE DSU TSD 10).

b) In its clarification letter, the EAG requested the company to explore different model types, i.e., Adjusted Limited Dependent Variable Mixture Model (ALDVMM) and Censored Least Absolute Deviations (CLAD) model. In their response, the company stated that this was not feasible in the time given to respond and provided additional arguments why re-estimating the mapping function would be unnecessary. The EAG also requested detailed responses to all aspects/considerations mentioned in Tables 1, 2 and 3 of the ISPOR Good Practices for mapping studies, but instead, the company highlighted only two aspects (related to the model choice and the mean predicted versus mean observed utility) that were considered important. The EAG believes that exploring different model types and providing details of *all aspects/considerations* of The Professional Society for Health Economics and Outcomes Research (ISPOR) Good Practices for mapping studies contributes to addressing the uncertainty surrounding the mapping of EQ-5D utilities from ISI© scores and hence prefers its request to be fulfilled.

4.2.9 Resources and costs

The cost categories included in the model were treatment costs and medical costs (GP, emergency room attendances, inpatient care). No costs associated with managing AEs were included in the model. Indirect costs associated with productivity were incorporated into a scenario analysis.

Unit costs were derived from the Personal Social Services Research Unit (PSSRU) 2021[33] and the NHS National Cost Collection (NCC) data for 2019/2020.[34]

4.2.9.1 Resource use and costs data identified in the review

An SLR was conducted (CS Appendix I) to identify the relevant studies evaluating costs and healthcare resource utilisation (HCRU) for patients with chronic insomnia disorder. Two of the six included studies included data from the UK as part of a larger multinational analysis.[35, 36] One study investigated the association between chronic insomnia disorder, indirect costs, and HCRU. Resource use and cost data identified in the SLR were not used in the economic model.

4.2.9.2 Treatment costs

The model only incorporates a daridorexant dosage of 50 mg. Therefore, costs are only reported for 50 mg. The treatment cost associated with daridorexant is given as [redacted] per day, giving an annual cost of [redacted] per patient. Conditional dropout rates from studies 301 and 303 were used as a proxy for discontinuation rates. The annual cost was adjusted for discontinuation to give an annual cost of [redacted]. For patients receiving "no-treatment", no treatment costs were incorporated.

4.2.9.3 Health state costs

The CS did not consider different health states. In response to clarification questionnaire B16, the company clarify that, instead of using differing health states, the mapping function of the ISI© total score and healthcare resource use and costs was utilised.

The association between direct healthcare resource use (related to GP visits, emergency room attendances, and inpatient care) and ISI© score were calculated from the NHWS data. This was done using a generalised linear model with a negative binomial distribution family and a log link. Costs were calculated by combining the estimated resource use with unit costs from the PSSRU 2021 (GP visits) and NHS England 2019/2020 costs (emergency room and inpatient costs), inflated to 2021 costs using the CPI index 06: Health.[37]

Table 4.8 outlines the total costs, per patient, included in the model for both the treatment and comparator for the 12-month time horizon.

Table 4.8: 12-month costs per patient

	No treatment	Daridorexant 50 mg
Tx costs	£0	[redacted]
GP costs	£323	£310
ER costs	£90	£83
IP costs	£211	£202
Total costs	£624	[redacted]

Based on CS Economic model
CS = company submission; Tx costs = treatment costs; GP = general practitioner; ER = emergency room; IP = inpatient care

4.2.9.4 Event costs

No costs associated with the management of AEs were included in the model for the intervention or comparator.

4.2.9.5 Productivity losses (used in scenario analyses only)

The company included a scenario analysis which explored the impact of productivity losses on the economic model. The CS estimated productivity losses from chronic insomnia disorder in two ways: 1) directly from the Sheehan Disability Scale (SDS) included in the clinical programme and 2) indirectly from the Work Productivity and Activity Impairment (WPAI) questionnaire mapped to ISI© in the NHWS dataset (calculating costs associated with productivity losses as a function of ISI© score).

To directly estimate the costs associated with productivity losses, SDS data from Study 301 and Study 303 were used. Total costs associated with productivity losses were calculated separately for absenteeism and presenteeism, and then combined. For calculating productivity losses 255 working days per year, 7.5 working hours per day, and the median annual wage rate for 2021 were assumed. The level of absenteeism was derived directly from item 4 (days lost) of the SDS. To derive the whole time equivalent (WTE) days lost due to presenteeism, the company weighted the number of days patients reported being underproductive (SDS item 5) by the level of unproductiveness (SDS item 1: extent symptoms have disrupted work/schoolwork). Over the 12-month time horizon, productivity savings, for those treated with daridorexant for the full year, are [redacted] ([redacted] after dropout adjustment), compared with the placebo. The EAG could largely replicate the cost calculations, with some minor discrepancies.

To indirectly estimate the productivity losses, the NHWS dataset was used which included administration of the WPAI questionnaire. The WPAI provided the hours missed due to health problems, hours actually worked, and the degree to which health affected productivity whilst working, which were used to estimate percentages for absenteeism and presenteeism. The percentages were used as the explanatory variable in a log-link generalised linear model with ISI© score as an explanatory variable. The percentage of absenteeism and presenteeism, as a function of ISI©, were costed utilising the median annual wage rate (£25,971). Over the 12-month time horizon, productivity savings for those treated with daridorexant for the full year, are [redacted] ([redacted] after dropout adjustment), compared with the placebo.

EAG comment:

- The main concerns of the EAG relate to: a) incorporated cost categories; and b) estimating productivity losses.

  a) When determining the non-treatment costs used in the economic model, the company only included resource use and costs in three categories: GP visits, emergency room visits, and inpatient care. Clarification question B16d requested justification for these three resource categories being the only relevant costs (in addition to direct treatment costs). In response, the company stated that these were the only costs available in the NHWS data set. The company acknowledges that other cost categories (e.g., concomitant medication use), however the company suggest the approach used was conservative. This was justified in the company response by stating that, as disease costs are background costs which an improved ISI© score will offset, all missing cost categories act against the treatment. The EAG would prefer that the company modelled all cost categories relevant to the NHS and PSS perspective, so as to support their justification and provide a more accurate reflection of the relevant costs associated with treatment implementation.

  b) The company included a scenario analysis incorporating the impact of including costs associated with productivity losses into the economic model. This was done using two methods: directly from SDS data, and indirectly through mapping WPAI data to the ISI©. The results for both methods reflected a reduction in costs associated with productivity losses for the treatment, compared with the no-treatment group. However, large discrepancies existed between the two methods (£[redacted] difference between savings associated with the treatment compared with no-treatment). It is unclear to the EAG which method would be a more accurate representation of costs associated with productivity losses. When directly estimating the productivity costs from SDS data, the company used SDS item 1 (extent symptoms have disrupted work/schoolwork) as a proxy for the level of unproductiveness on days at work. Item 1 does not specify that respondents should only consider time spent at work, and to not consider the impact of their symptoms on absenteeism. As such, it remains unclear whether item 1 is a plausible proxy for the level of unproductiveness. For the indirect method, the WPAI data is mapped to ISI© scores, and subsequently mapped to resource use/costs. The company does not discuss the uncertainty associated with using this approach. Uncertainty surrounding both methods in addition to being unable to determine which method is likely to be superior, culminate to the EAG being unable to suggest it is satisfied with the reported savings associated with productivity losses for the treatment, compared with no-treatment.

4.2.10 Severity

The initial CS [1] stated that the Severity Section (B.3.6) is *"not relevant to this submission"*. In response to clarification question B23 the company stated: *"The QALY shortfall does not justify an additional severity weighting and we therefore assume a QALY weight of 1 will apply"*. [3]

EAG comment:

- The EAG believes assuming a QALY weight of 1 is reasonable.

4.2.11 Uncertainty

The CS stated that the Uncertainty Section (B.3.7) is "not relevant to this submission". In response to clarification question B24 the company stated that this Section is only relevant "for new technologies where there may be significant gaps in the evidence base".

EAG comment:

- The EAG disagrees with the company. There are significant uncertainties in the CS, as described in the previous Sections, that would warrant uncertainty analyses to show the impact on the cost effectiveness results.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base case cost effectiveness results (probabilistic) indicated that daridorexant is both more effective (incremental QALYs 0.024) and more costly (additional costs of [redacted]) than no-treatment amounting to an ICER of £[redacted] per QALY gained (Table 5.2). Moreover, the 95% percentiles for the probabilistic incremental costs and QALYs were (£[redacted]) and (0.015 – 0.034) respectively. The probability of daridorexant being cost-effective at threshold of £30,000 per QALY gained compared to no-treatment is [redacted].

Table 5.1: Company deterministic base case results, adjusted for dropout

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER	iNHB (£20,000)	iNHB (£30,000)
Daridorexant	[redacted]	0.725					
No-treatment	£624	0.691	[redacted]	0.024*	[redacted]	[redacted]	[redacted]

Based on CS Table 57 [1]
*Adjusted for dropout
CS = company submission; ICER = incremental cost effectiveness ratio, iNHB = incremental net health benefit; QALY = quality adjusted life year

Table 5.2: Company probabilistic base case results, adjusted for dropout

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER	iNHB (£20,000)	iNHB (£30,000)
Daridorexant	NR	NR					
No-treatment	NR	NR	[redacted]	0.024*	[redacted]	[redacted]	[redacted]

Based on CS Table 58 [1]
*Adjusted for dropout
CS = company submission; ICER = incremental cost effectiveness ratio, iNHB = incremental net health benefit; NR = not reported; QALY = quality adjusted life year

Overall, the technology is modelled to affect QALYs by:

- The ISI© scores of Study 301 and Study 303
- The ISI© score to EQ-5D mapping algorithm

Overall, the technology is modelled to affect costs by:

- Treatment costs
- Health care costs
- Productivity loss (in scenario analyses)

5.2 Company's sensitivity analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSAs), deterministic sensitivity analyses (DSAs) as well as scenario analyses. The parameters that have the greatest effect on the ICER (based on the company's sensitivity analyses) are:

- The ISI© scores of studies 301 and 303
- The parameters of the mapping algorithm of the ISI© scores to European Quality of Life-5 Dimensions (EQ-5D)

CS scenarios that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- Inclusion of indirect costs [redacted] per QALY gained)
- Optimistic scenario [redacted] per QALY gained)
- Pessimistic scenario (£[redacted] per QALY gained)

EAG comment:

- No comment.

5.3 Model validation and face validity check

5.3.1 Face validity assessment

According to the CS, the underlying concept of the analyses was presented to several clinical experts, health technology assessment (HTA) experts and NICE (during the decision problem meeting) to assess the face validity of the economic model.

5.3.2 Technical verification

Technical verification was conducted by Avalon Health Economics. This included testing of the model programming and replicating the results by individuals who were not involved in the development of the model.

5.3.3 Comparisons with other technology appraisals

The company did not provide any cross comparisons with other relevant technology appraisals in the CS section B.3.14

5.3.4 Comparison with external data used to develop the economic model

The company did not provide comparisons with external data used to develop the economic model in the CS section B.3.14.

5.3.5 Comparison with external data not used to develop the economic model

The company did not provide a comparison with external data not used to develop the economic mode in the CS section B.3.14.

EAG comment:

- The main concerns of the EAG relate to a) the few face validity and technical validity assessment details in the initial submission, b) the lack of cross validation with other relevant technology appraisals and the lack of external validation of the model.

    a) The EAG was concerned about the few face validation and technical validation details that were provided in the initial submission. In their response to clarification question B20, the company restated their validation process in reference to the TECH-VER checklist[38], including the pre-analysis assessment of completeness and consistency, calculation consistency between the model and the description and values, the correctness of the model implementation, event and results in calculation and the validation of the model uncertainty analysis and scenario analyses. Despite that the company did not provide a fully filled checklist, the EAG is satisfied with the face validity and technical validity evidence provided by the company.

b) The company did not provide a cross validation to other relevant technology appraisals in its initial submission. Neither it provided any information regarding external model validation. Upon request, the company stated that the models presented in other appraisals (i.e., TA77[39] and MTG70[40]) are not full cost-per-QALY models for insomnia and were therefore not directly comparable to the model results and hence no cross-validation could be done. Furthermore, the company did not provide any external validation for the data used to develop the economic model. The company updated the model to include the use of the mapping function by Gu et al[41]. as an alternative for the NHWS algorithm, which increased the ICER from [redacted] to [redacted]. The EAG acknowledges that further cross validation to other relevant appraisals may not be feasible but prefers further external validation of the economic model.

6. EVIDENCE ASSESSMENT GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020[42]:

- Transparency (e.g., lack of clarity in presentation, description, or justification)
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case)
- Imprecision (e.g., particularly wide confidence intervals, small sample sizes, or immaturity of data)
- Bias & indirectness (e.g., there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered)
- Unavailability (e.g., lack of data or insight)

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/ or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the cost effectiveness, whether it is reflected in the EAG base case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this EAG report, the EAG defined a new base case. This base case included multiple adjustments to the original base case presented in the previous sections. These adjustments made by the EAG form the EAG base case and were subdivided into three categories (derived from Kaltenthaler 2016[43]):

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV) (correcting the model where the EAG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ) (amending the model where the EAG considers that reasonable alternative assumptions are preferred)

6.1.1 EAG base case

Adjustments made by the EAG, to derive the EAG base case (using the CS base case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the EAG base case. The 'FE' adjustments were combined and the other EAG analyses were performed also incorporating these 'FE' adjustments given the EAG considered that the 'FE' adjustments corrected unequivocally wrong issues.

6.1.1.1 Fixing errors

1. Exclusion of dropout in the no treatment arm. (Section 4.2.6)

Although the company stated in the CS that patients could not dropout from no-treatment, this was not implemented in the economic model as such and hence corrected by the EAG.

6.1.1.2 Fixing violations

No violations were identified by the EAG.

6.1.1.3 Matters of judgement

2. Patients dropping out from daridorexant were assumed to have no impact on HRQoL and costs (Section 4.2.8 and 4.2.9).

After fixing the assumption that patients in the no-treatment arm could not dropout from no treatment, in the economic model, QALYs and costs were calculated based on the proportion of patients that were not dropped out, i.e., patients who dropped out were assumed to have no impact on HRQoL and costs. The EAG assumed the proportion of patients dropping out of the daridorexant arm in its base case to revert to the HRQoL and costs as assigned to the placebo arm.

3. The placebo effect in the no treatment arm was only applied in the first 3 months (e.g., based on Study 301 only) (Section 4.2.6).

The company did not include the placebo effect observed in Study 303 (the study that informs months 3-12 in the economic model), arguing that Study 303 presented evidence of selective attrition. The EAG disagreed and applied the placebo effect in the no treatment arm to the whole (1 year) time horizon of the model.

6.1.2 EAG exploratory scenario analyses

The EAG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the EAG base case.

6.1.2.1 Exploratory scenario analyses

4. Informing dropout rates based on Study 303 (Section 4.2.6).

In the company's base case, dropout rates for months 1 and 3 were informed based on Study 301 whereas the dropout rates in the remaining months were informed based on Study 303. The EAG explored a scenario analysis informing dropout rates for all time point based on Study 303.

6.1.3 EAG subgroup analyses

No subgroup analyses were performed by the EAG.

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1)

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG base case[b]	Required additional evidence or analyses
Patients with mental health problems were excluded from the studies informing the economic model, decreasing the generalizability of the underlying evidence to the decision problem.	4.2.3	Bias and indirectness	New evidence for patients with comorbid mental health problems receiving different treatments.	+/-	No	New evidence for patients with comorbid mental health problems receiving different treatments.
The company only included no-treatment as a comparator to daridorexant in the health economic model.	4.2.4	Bias and indirectness	Inclusion of all relevant comparators.	+/-	No	Inclusion of all relevant comparators.
The 25 mg dosage of daridorexant (part of the anticipated market authorisation) was not included in the economic model.	4.2.4	Bias and indirectness	New evidence for the 25 mg dosage subgroup. Scenario analysis using only the 25 mg population in Study 301 and Study 303.	+/-	No	New evidence for the 25 mg dosage subgroup. Scenario analysis using only the 25 mg population in Study 301 and Study 303.
In contrast to what was stated in the CS, the no-treatment arm dropout rates observed in Study 301 and Study 303 were also applied to the no-treatment arm in the economic model.	4.2.6	Methods	Applying no dropout rates for the no-treatment arm.		Yes	N/A
Placebo was only adjusted for the first 3 months in the no-treatment arm, but not for the remaining 40 weeks. This could	4.2.6	Methods	The pessimistic scenario applying the placebo effect observed in studies	+	Yes	N/A

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG base case[b]	Required additional evidence or analyses
have biased the comparison in favour of the intervention.			301 and 303 to both arms.			
The company did not include AEs from studies 301 and 303 in their economic model, assuming that these are minor AEs and would not be expected to have consequences on resource use or HRQoL.	4.2.7	Methods	An updated cost effectiveness model and scenario analyses incorporating all AEs from studies 301 and 303.	+/-	No	An updated cost effectiveness model and scenario analyses incorporating all AEs from studies 301 and 303.
There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.	4.2.8	Bias and indirectness	Scenario analysis incorporating a re-estimated mapping function in line with ISPOR Good Practices for mapping studies and including relevant covariates. Scenario analyses exploring ALDVMM and CLAD models. Detailed responses to all aspects/considerations mentioned in Tables 1, 2 and 3 of the ISPOR Good Practices for mapping studies.	+/-	No	Scenario analysis incorporating a re-estimated mapping function in line with ISPOR Good Practices for mapping studies and including relevant covariates. Scenario analyses exploring ALDVMM and CLAD models. Detailed responses to all aspects/considerations mentioned in Tables 1, 2 and 3 of the ISPOR Good Practices for mapping studies.

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG base case[b]	Required additional evidence or analyses
Not all potentially relevant costs were included in the economic model	4.2.9	Methods	Identification and inclusion of all additional cost categories, relevant to the NHS/PSS perspective, into the economic model.	+/-	No	Identification and inclusion of all additional cost categories, relevant to the NHS/PSS perspective, into the economic model.

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the EAG and '+' indicates that the EAG believes this issue likely induces bias in favour of the intervention versus at least one comparator; [b] Explored

AEs = adverse events; ALDVMM = Adjusted Limited Dependent Variable Mixture Model; CLAD = Censored Least Absolute Deviations; CS = company submission; EAG = Evidence Review Group; EQ-5D = European Quality of Life-5 Dimensions; FE = Fixing errors; FV = fixing violations; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; ISI© = Insomnia Severity Index; ISPOR = The Professional Society for Health Economics and Outcomes Research; MJ = matters of judgement; N/A = not applicable; NHS = National Health Service; PSS = Personal Social Services

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the EAG

In Section 6.1 the EAG base case was presented, which was based on various changes compared to the company base case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3. These are all conditional on the EAG base case. The analyses numbers in Table 6.2 and Table 6.3 correspond to the numbers reported in Section 6.1.

Table 6.2: Deterministic EAG base case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case (without dropout)					
Daridorexant	[redacted]	0.725			
No-treatment	£624	0.691	[redacted]	0.034	[redacted]
CS base case (with dropout)					
Daridorexant	[redacted]	0.543			
No-treatment	£471	0.519	[redacted]	0.024	[redacted]
Fixing errors (1 - No dropout for the no-treatment group)					
Daridorexant	[redacted]	0.543			
No-treatment	£624	0.691	[redacted]	-0.148	Daridorexant dominated by no-treatment
Matter of judgment (2 - Adjustment utility and costs for dropout patients)*					
Daridorexant	[redacted]	0.715			
No-treatment	£624	0.691	[redacted]	0.024	[redacted]
Matter of judgment (3 - Removing the company's placebo adjustment)*					
Daridorexant	[redacted]	0.543			
No-treatment	£614	0.703	[redacted]	-0.160	Daridorexant dominated by no-treatment
EAG base case					
Daridorexant	[redacted]	0.720			
No-treatment	£614	0.703	[redacted]	0.017	[redacted]

*Conditional of (1-No dropout for the no-treatment group)
CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Table 6.3: Deterministic scenario analyses (conditional on EAG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
EAG base case					
Daridorexant	[redacted]	0.720			
No-treatment	£614	0.703	[redacted]	0.017	[redacted]

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Scenario analysis (4 - Dropout rates based on Study 303 for daridorexant)					
Daridorexant	[redacted]	0.717			
No-treatment	£614	0.703	[redacted]	0.014	[redacted]
EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year					

Table 6.4: Probabilistic EAG base case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case (without dropout)*					
Daridorexant	[redacted]	0.724			
No-treatment	£637	0.691	[redacted]	0.034	[redacted]
CS base case (with dropout)*					
Daridorexant	[redacted]	0.543			
No-treatment	£478	0.518	[redacted]	0.024	[redacted]
EAG base case					
Daridorexant	[redacted]	0.720			
No-treatment	£622	0.703	[redacted]	0.017	[redacted]

*These results are slightly different from the ones stated in the CS, due to:

- The Excel model code
- The EAG has calculated the ICER from the total costs and QALYs from the PSA, not the incremental results from those

CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

Table 6.5: Probabilistic EAG scenario analyses

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
EAG base case					
Daridorexant	[redacted]	0.720			
No-treatment	£622	0.703	[redacted]	0.017	[redacted]
Scenario analysis (4 - Dropout rates based on Study 303 for daridorexant)					
Daridorexant	[redacted]	0.720			
No-treatment	£621	0.703	[redacted]	0.017	[redacted]
EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year					

6.3 EAG's preferred assumptions

The estimated EAG base case ICER (probabilistic), based on the EAG preferred assumptions highlighted in Section 5.1, was £36,562 per QALY gained. The most influential adjustment was related to the company's placebo adjustment. The ICER increased in the scenario analysis with alternative assumptions regarding dropout rates.

6.4 Conclusions of the cost effectiveness section

The CS[1] and response to clarification[3] provided sufficient details for the EAG to appraise the literature searches conducted to identify economic, HRQoL and cost data on chronic insomnia disorder. Searches were conducted in April 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases were searched. Overall, the EAG has no major concerns about the literature searches conducted, although a broader approach to conference searching may have retrieved additional studies.

The company's cost effectiveness model partly complied with the NICE reference case. Deviations from the NICE reference case related to the synthesis of evidence on health effects as no review was used to identify relevant mapping functions or sources that could potentially be used to develop a mapping function. Moreover, it was unclear 1) whether the mapping function (to estimate EQ-5D utilities) was based on the UK tariff and 2) whether all relevant costs and effects were captured within the 12-month time horizon. The most prominent issues highlighted by the EAG were 1) the generalisability of the treatment effect to the anticipated treatment population in the UK; 2) not including all comparators mentioned in the scope; 3) not considering the 25 mg dosage of daridorexant (part of the anticipated market authorisation)); 4) justification for the use of the ISI© score for the estimation of treatment effectiveness; 5) the company's placebo adjustment and 6) uncertainties related to the mapping function.

Firstly, the exclusion of patients with mental health problems and the inclusion of patients receiving CBT-I in studies 301 and 303 results in uncertainty surrounding the generalisability of the treatment effect to the anticipated treatment population. Secondly, although a variety of pharmaceuticals and therapies are available for the treatment of insomnia, the company only included 'no-treatment' as a comparator to daridorexant in the health economic model. Relevant comparators such as sleep hygiene advice, CBT-I, non-benzodiazepine hypnotic medication, zolpidem, zopiclone, benzodiazepines and melatonin were excluded. Thirdly, the company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation and according to the company relevant where there is co-administration of moderate CYP3A4 inhibitors. Fourthly, treatment effectiveness in the economic analyses was based on ISI© scores, an exploratory trial outcome, from studies 301 and 303. Other clinical primary (i.e., WASO and LPS), and secondary (i.e., sTST and IDSIQ) efficacy endpoints were collected in the trials but were not used to inform treatment effectiveness. The EAG considers that the use of ISI© scores was not sufficiently justified. Fifthly, the company applied the placebo effect in the no treatment arm only for the first three months. The EAG disagrees with this approach and applied a placebo correction in the no treatment arm to the whole (1 year) time horizon of the model. Finally, there were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.

The CS base case probabilistic and deterministic ICERs (with dropout adjustment) were [redacted] and [redacted] per QALY gained, respectively. The EAG base case probabilistic and deterministic ICERs, based on the EAG preferred assumptions highlighted in Section 6.1, were [redacted] and [redacted] per QALY gained. The most influential adjustment was related to the company's placebo correction for the no treatment arm. The ICER [redacted] by ~£[redacted] in the scenario assuming alternative dropout rates.

Remaining uncertainty about the effectiveness and relative effectiveness of daridorexant can be at least partly resolved by the company by conducting further analyses (as highlighted in Table 6.1) and

providing further justification regarding the appropriateness of the mapping function. Moreover, the current assessment does not provide an appropriate estimation of the comparators listed in the scope.

7. END-OF-LIFE

The company does not claim that the intervention meets the NICE end-of-life criteria.

EAG comment:

- The EAG agrees that the intervention does not meet the NICE end-of-life criteria.

8. REFERENCES

[1] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*: Idorsia Pharmaceuticals Ltd., 2022. 144p.

[2] National Institute for Health and Care Excellence. *Daridorexant for treating insomnia disorder [ID3774]: Clarification questions*. London: NICE, 2022. 32p.

[3] National Institute for Health and Care Excellence. *Daridorexant for treating insomnia disorder [ID3774]: Response to request for clarification from the ERG*. London: National Institute for Health and Care Excellence, 2022. 101p.

[4] National Institute for Health and Care Excellence. *Daridorexant for treating insomnia disorder. Pre-invite scope [Internet]*. London: NICE, 2021 [accessed 5.7.22]. 3p. Available from: https://www.nice.org.uk/guidance/gid-ta10888/documents/final-scope

[5] Morin CM, Bélanger L, LeBlanc M, Ivers H, Savard J, Espie CA, et al. The natural history of insomnia: a population-based 3-year longitudinal study. *Arch Intern Med* 2009;169(5):447-53.

[6] Morphy H, Dunn KM, Lewis M, Boardman HF, Croft PR. Epidemiology of insomnia: a longitudinal study in a UK population. *Sleep* 2007;30(3):274-80.

[7] Idorsia Pharmaceuticals Ltd. *Idorsia receives US FDA approval of QUVIVIQ (daridorexant) 25 and 50 mg for the treatment of adults with insomnia. Media Release, January 10, 2022 [Internet]*: Idorsia Pharmaceuticals Ltd., 2022 [accessed 19.9.22] Available from: https://www.idorsia.us/documents/us/media-releases/2201_us-fda-approval-quviviq-announcement.pdf

[8] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[9] Canadian Agency for Drugs and Technologies in Health. *PRESS - Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 13.7.22] Available from: https://www.cadth.ca/resources/finding-evidence/press

[10] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 13.7.22]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[11] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix D: Identification, selection and synthesis of clinical evidence*: Idorsia Pharmaceuticals Ltd., 2022. 84p.

[12] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 13.7.22] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[13] Golder S, Peryer G, Loke YK. Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *J Clin Epidemiol* 2019;113:36-43.

[14] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 4.3.22] Available from: https://training.cochrane.org/handbook

[15] Idorsia Pharmaceuticals Ltd. *ID-078A301: Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder. [CSR-301]*: Idorsia Pharmaceuticals Ltd., 2020. 339p.

[16] Idorsia Pharmaceuticals Ltd. *ID-078A302: Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder [CSR-302]*: Idorsia Pharmaceuticals Ltd., 2020. 356p.

[17] Dauvilliers Y, Zammit G, Fietze I, Mayleben D, Kinter DS, Pain S, et al. Daridorexant, a new dual orexin receptor antagonist to treat insomnia disorder. *Ann Neurol* 2020;17.

[18] Idorsia Pharmaceuticals Ltd. *ID-078A303: Multi-center, double-blind, parallel-group, randomized, placebo-controlled, three doses, 40-week extension to studies ID-078A301 and ID-078A302 to assess the long term safety and tolerability of ACT-541468 in adult and elderly subjects with insomnia disorder. Final CSR [Data on file]*: Idorsia Pharmaceuticals Ltd., 2021. 328p.

[19] Office for National Statistics. Population of England and Wales [Internet]. 2020 [accessed 19.9.22]. Available from: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/

[20] Fernandez-Mendoza J, Bourchtein E, Calhoun S, Puzino K, Snyder CK, He F, et al. Natural history of insomnia symptoms in the transition from childhood to adolescence: population rates, health disparities, and risk factors. *Sleep* 2021;44(3):zsaa187.

[21] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix M: Clinical effectiveness*: Idorsia Pharmaceuticals Ltd., 2022. 28p.

[22] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix F: Adverse reactions*: Idorsia Pharmaceuticals Ltd., 2022. 24p.

[23] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix E: Subgroup analysis*: Idorsia Pharmaceuticals Ltd., 2022. 21p.

[24] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix G: Published cost-effectiveness studies*: Idorsia Pharmaceuticals Ltd., 2022. 49p.

[25] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix H: Health-related quality of life studies*: Idorsia Pharmaceuticals Ltd., 2022. 61p.

[26] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix I: Cost and healthcare resource identification, measurement and valuation*: Idorsia Pharmaceuticals Ltd., 2022. 37p.

[27] Reddy MS, Chakrabarty A. 'Comorbid' insomnia. *Indian J Psychol Med* 2011;33(1):1-4.

[28] Wichniak A, Wierzbicka A, Walęcka M, Jernajczyk W. Effects of antidepressants on sleep. *Curr Psychiatry Rep* 2017;19(9):63.

[29] National Institute for Health and Care Excellence. NICE Clinical Knowledge Summaries: Insomnia [Internet]. 2022 [accessed May 2022]. Available from: https://cks.nice.org.uk/topics/insomnia/

[30] Idorsia Pharmaceuticals Ltd. *Daridorexant for treating insomnia disorder [ID3774]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document A - Company evidence submission summary for committee*: Idorsia Pharmaceuticals Ltd., 2022. 28p.

[31] Longworth L, Rowen D. *NICE DSU Technical Support Document 10: The use of mapping methods to estimate health state utility values*. Sheffield: Decision Support Unit, ScHARR, 2011. 31p. Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD10-mapping-FINAL.pdf

[32] Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010;11(2):215-25.

[33] Jones KC, Burns A. *Unit Costs of Health and Social Care 2021*. Kent: Personal Social Services Research Unit, 2021. 185p.

[34] NHS. National Cost Collection for the NHS [Internet]. [accessed 8.2.22]. Available from: https://www.england.nhs.uk/national-cost-collection/

[35] DiBonaventura M, Richard L, Kumar M, Forsythe A, Flores NM, Moline M. The association between insomnia and insomnia treatment side effects on health status, work productivity, and healthcare resource use. *PLoS One* 2015;10(10):e0137117.

[36] Matos JE, Chalet FX, Vaillant C, Roberts G. POSC430 Unmet needs for insomnia patients - economic burden in Europe 2022;25(1):S277.

[37] Office for National Statistics. CPI weights 06: Health [Internet]. 2022 [accessed 6.9.22]. Available from: https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/chzw/mm23

[38] Buyukkaramikli NC, Rutten-van Molken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics* 2019;37(11):1391-1408.

[39] National Institute for Health and Care Excellence. *Guidance on the use of zaleplon, zolpidem and zopiclone for the short-term management of insomnia [TA77]*. London: NICE, 2004 [accessed May 2022] Available from: https://www.nice.org.uk/guidance/ta77

[40] National Institute for Health and Care Excellence. *Sleepio to treat insomnia and insomnia symptoms. Medical technologies guidance [MTG70]*. London: NICE, 2022 [accessed May 2022] Available from: https://www.nice.org.uk/guidance/mtg70

[41] Gu NY, Botteman MF, Ji X, Bell CF, Carter JA, van Hout B. Mapping of the Insomnia Severity Index and other sleep measures to EuroQol EQ-5D health state utilities. *Health Qual Life Outcomes* 2011;9:119.

[42] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-216.

[43] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

Single Technology Appraisal

Daridorexant for treating insomnia [ID3774]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on 29 September 2022** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and separately highlight information that is submitted as 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

Issue 1 Comparators of daridorexant as specified in NICE decision problem

Description of problem	Description of proposed amendment	Justification for amendment	EAG
Key issue 3 and key issue 9 in sections 2.3 and 4.2.4 lack context and do not reflect the discussion and agreement prior to final scope and in the subsequent DPM. The discussion in the CS regarding comparators was based on these agreements. As it is written, the EAG report suggests otherwise (e.g., "*The CS therefore fails to present data relating to the decision problem*")*,* which is factually inaccurate.	The company proposes the EAG to review key issue 3 and key issue 9, considering the discussions with NICE during the scoping meeting and DPM. Further information is provided below for context: The position regarding active comparators should reflect the scoping and DPM decisions. To acknowledge this, the text from scoping actions should be referred to: "*It was noted by clinical experts and patient group representatives that access to CBT-I varies geographically. They also noted that CBT-I is not always offered. At the workshop, attendees also discussed pharmacological treatments for insomnia. They explained that none of the currently approved pharmacological treatments are recommended for long term use. Daridorexant is expected to be used to treat insomnia disorder, where symptoms last for more than 3 months per clinical trial. Therefore, the attendees agreed that none of the comparators listed in the draft scope are relevant. The scope has been updated to remove the comparators*" The following comparators were listed in the draft scope:	It is a fundamental criticism from the EAG report that the CS does not answer the decision problem. Regarding comparators, this lacks context to previous discussions and agreements. It also fails to acknowledge two critical facts unique to this CS: 1. The long-term use of existing pharmacological therapies is against all current clinical guidance and expert opinion for the treatment of insomnia disorder. The use of these therapies, beyond their recommended licensed duration, is identified as a significant clinical problem with specific medicines management guidance published by NICE in 2019, 'Benzodiazepine / Hypnotics de-prescribing' with further guidance published in 2022, 'Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults' to reduce	No factual inaccuracy. The EAG responded to the information given in the CS.

	• Established clinical management (including sleep hygiene and CBT-I) • Zolpidem and zopiclone • Melatonin (for those aged 55 and over) • Benzodiazepines (for example nitrazepam, loprazolam, lormetazepam, temazepam). The final scope was updated after the scoping meeting, where all comparators were removed. Whilst established clinical management remained, CBT-I was removed: • Established clinical management (including sleep hygiene advice) without daridorexant It was also discussed in the DPM that the population with insomnia in the clinical studies and in the real world would already have completed basic sleep hygiene measures prior to pharmacological therapy. The relevant section from the DPM form used in the meeting is replicated below. It is noted that the EAG was not present in the scoping meeting and a separate process letter has been sent to NICE on this point.	prescribing and withdraw patients. It cannot be considered established clinical management as comparison and this was agreed in scoping (see text opposite) 2. CBT-I is the established gold standard first line treatment as recommended in both European and UK treatment guidelines. However, in the same survey of 1,001 UK GPs, only around 20% of the GPs reported that NHS-funded CBT-I was available to them; however, CBT-I access was limited by capacity and waiting times. The primary positioning of daridorexant is after CBT-I, however due to the above-mentioned reasons, many patients do not have access to CBT-I, or may refuse CBT-I, and prescribers reluctantly resort to non-licensed and non-recommended pharmacotherapies. It is this real-world CBT-I issue that drives the daridorexant secondary positioning as an alternative in these situations, where CBT-I is not an option, as opposed to seeking to compare with CBT-I. If	

Description of problem	Description of proposed amendment	Justification for amendment	EAG
	<table><tr><td>Comparator(s)</td><td>None. The current pharmacological treatments for chronic insomnia disorder are not approved or recommended for long-term use as stated by the experts present at the initial Scoping meeting and reflected in the NICE Clinical Knowledge Summary for managing long term insomnia. Consequently, they were removed in the pre-invitation scope.</td></tr></table> On this basis it was agreed placebo comparison was appropriate, albeit there is no placebo in real world practice, and this would likely understate the benefits of daridorexant. This is also reflected in data obtained from a study in 1,001 UK GPs that reported that a majority tend to start with non-pharmacological therapies such as sleep hygiene advice, OTC treatment, face-to-face or digital CBT-I.	it simplifies this amendment, the company is satisfied with a second line position, after CBT-I has been offered to patients. To give additional context the EAG notes that they are “surprised” that 87.9% of insomnia disorder patients in the clinical trials either did not know about or were never offered CBT-I. This reflects the real-world issue we are seeking to address, ie, that availability and funding for CBT-I is significantly limited and leaves GPs with no option other than to use non-recommended therapies in some patients. In view of these facts, the company <u>did not</u> (as opposed to, it failed to) perform an indirect treatment comparison with CBT-I to be consistent with the NICE scope.	

Issue 2 Application of dropout rates to both daridorexant and no treatment groups in the economic model

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 1.1, Page 14, Key issue 11 – The EAG stated that "*…in the economic model provided by the company, the dropout rates observed in studies 301 and 303 for the daridorexant arm were applied to both daridorexant and no-treatment groups. This contradicts the statement made by the company.*" In the EAG's corrected analysis, (Table 6.2, page 117), the corrected model showed that daridorexant was dominated by no treatment. This is factually inaccurate and likely reflects a misunderstanding of what was actually done in the model formula and calculation. However,	The company proposes the EAG to review its corrected analyses and key issue 11.	The EAG appears to be saying the following: • They agreed that it makes no sense that no treatment can have dropout as there is nothing to drop out from • They believe the company has incorrectly included drop out from "no treatment", (which was not the case). • When they correct this perceived error treatment with daridorexant becomes more expensive and less effective than no treatment. It was not possible to assess the corrected analysis under question as it was not made available to the company, instead the company presents the below detailed explanation of how dropout was applied in the health economic model. Considering quality of life, Q, but without loss of generality as this could apply to any measure that is compared between the treatment arms. Further, without any dropout adjustment, the	No factual inaccuracy. Due to a lack of model transparency and limited justification by the company on how drop-out was modelled, the ERG considered the company's approach of modelling drop-out was incorrect. Based on the clarification provided by the company as part of the FAC, the EAG would like to highlight that it implemented drop-out in line with the company's approach, but in two steps: 1) adjusting for drop-out in the no-treatment group (with 0% drop-out) and the treated group (drop-out rates from the trials), 2) applying the utility of patients with no-treatment to patients treated with daridorexant that dropped-out. The EAG considers this form of presenting the results more transparent than the company's approach. Hence, the EAG base-case assumptions remain the same.

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
this needs to be confirmed as the company did not receive the corrected model which showed that daridorexant was dominated by no treatment.		difference in quality of life at a given time point is defined as: $$\Delta Q = Q_D - Q_{NT}$$ where the subscripts *D* and *NT* refer to the daridorexant and no-treatment groups, respectively. In the <*Model* | *1 year*> worksheet of the economic model of the CS: • quality of life (and QALYs) for daridorexant comes from row 34 • quality of life (and QALYs) is estimated in row 17 for no treatment • difference between them (no dropout adjustment) comes from row 51. Dropout was adjusted in the model as follows: 1) Persistence, *p*, was defined as one minus dropout. So, *p* reflects those staying on treatment and (1 - *p*) reflects	

		those that drop out from active treatment. 2) It was assumed that those staying on treatment get the Q_D outcome and those dropping out get the Q_{NT} outcome. Therefore, dropout was adjusted by multiplying the unadjusted, ΔQ, by the persistence, p, to give the adjusted $\Delta Q'$ as: $$\Delta Q' = p\Delta Q = pQ_D - pQ_{NT}$$ Note this is the calculation in row 63 of the economic model of the CS. The expansion on the right-hand side of the equation may have led the EAG to assume that the model has 'adjusted both the daridorexant and no-treatment groups for dropout'. Next, the dropout corrected quality of life in the treatment arm is defined as: $$Q'_D = pQ_D + (1-p)Q_{NT}$$ that is, a weighted average of the quality of life on treatment and the quality of life for those who dropped out from treatment weighted by the respective persistence and dropout percentages. The dropout-corrected difference in quality of life (or QALYs) is defined as: $$\Delta Q' = Q'_D - Q_{NT}$$	

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
		While the above expression is consistent with the EAG's expectations (i.e., a treatment arm that is adjusted for dropout and a no treatment arm that is unadjusted), it is equivalent to the one implemented in the economic model of the CS as shown below. Substitute the definition of Q'_D into the above expression: $$\Delta Q' = pQ_D + (1-p)Q_{NT} - Q_{NT}$$ This can be expanded to: $$\Delta Q' = pQ_D + Q_{NT} - pQ_{NT} - Q_{NT}$$ and further simplified to: $$\Delta Q' = pQ_D - pQ_{NT}$$ This demonstrates that the model in the CS is correct, and it is factually incorrect to say the model needs correction. This gives the impression the company has been careless, and the model is not to be trusted.	

Issue 3 Placebo correction in the no treatment group of the health economic model

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
Table 1.1, Page 14, Key issue 12 – The EAG stated that "*For the company base case, the placebo effect was only included for the first-three months in the no-treatment arm, but not for the remaining 40 weeks.*". This is factually incorrect as a placebo effect was applied to the entire 12 months of the model and is clearly shown in the base-case model diagram.	The company proposes that the EAG reviews key issue 12 considering that a placebo effect was applied to the no treatment group based on study 301 and extrapolated to the entire duration of the model.	In the base case model (Fig 4.1 of EAG report, page 99), placebo effect was applied to the no treatment group using study 301 and the same rate was extrapolated to the entire duration of the model. As explained in Document B (B.3.3.2), it was assumed that the no treatment group would continue at the same ISI achieved by the end of study 301 (i.e., the 3rd month) due to evidence of selective attrition in study 303. This approach is supported by data from study 303. In study 303, rebound of ISI was observed during placebo run out in study 301 prior to study 303 entry. For subjects who continued from study 301 to study 303, similar placebo effect was observed in the first 3 months of study 303 (Figure 13, Document B, B.2.9.1). Therefore, it can be considered that study 303 is placebo corrected up to the first 3 months and selective attrition was applied thereafter in the economic model. The EAG did not find the company's argument against regression to the mean convincing. Having set out the clarification on placebo correction, for regression to the mean to be	No factual inaccuracy. It is a matter of judgement.

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
		explaining the selective attrition in study 303 after the four week run in before randomisation (followed by 12 weeks in study 301 and a further 14 weeks in study 303) is highly implausible. This study design largely eliminates the likelihood of regression to the mean.	

Issue 4 Latest data cut-off

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
Section 3.2.5, Page 66 – The EAG stated that "*In the clarification letter the company has been asked to confirm that the latest data cut-off was 22 July 2020 and provide newer data, if available, in an addendum. The company confirmed that it was **22 July 2022***.".	The statement should read: "In the clarification letter the company has been asked to confirm that the latest data cut-off was 22 July 2020 and provide newer data, if available, in an addendum. The company confirmed that it was **22 July 2020**."	The company confirms that the data cut-off was 22 July 2020 as stated in its clarification letter. This indicates that no new data are available after the cut-off date.	Report has been amended accordingly.

Issue 5 Change scores used in EAG's analyses of rebound insomnia

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
Section 3.2.5.5, Page 79 – The EAG stated that "*For WASO and LPS these results suggest, in contrast to the conclusions of the company, a significantly greater rebound effect for daridorexant than placebo, though this was not observed for sTST.*" This led to the EAG to conclude in Section 3.6, Page 89, that "*…daridorexant had a significantly greater risk of rebound insomnia in terms of WASO and LPS.*".	A negative change in WASO and LPS for the placebo group should be used in the EAG's analyses.	In the EAG's analyses of data on rebound insomnia in study 301, a positive change from baseline in WASO and LPS for the placebo group was used instead of a negative change. According to Appendix F of the CS (F.1.1.4), "Subjective total sleep time (sTST) were compared for the treatment and placebo group (WASO: –2.517 [52.355] for daridorexant 50 mg vs **-20.392** [45.776] for placebo and LPS: –15.035 [55.812] for daridorexant 50 mg vs **-27.820** [47.199] for placebo)." This will impact the mean difference and 95% confidence interval for WASO and LPS in the EAG's analyses, potentially resulting in a different conclusion on rebound insomnia.	Report has been amended accordingly with a new EAG analysis performed.

Issue 6 Name of health technology assessed in the CS

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
Section 3.2.5.7, Page 82 – The EAG stated that "*For IDSIQ, the lower body mass index (BMI) sub-group experienced a better response from* ***regorafenib*** *(relative to placebo) than the higher BMI sub-group.*" The technology under assessment should be daridorexant instead of regorafenib.	The statement should read "*For IDSIQ, the lower body mass index (BMI) sub-group experienced a better response from* ***daridorexant*** *(relative to placebo) than the higher BMI sub-group.*".	The correct health technology being assessed in this CS should be used, though this does not impact the assessment.	Report has been amended accordingly.

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
Section 1.2, Page 15	CIC not marked	• Inclusion of indirect costs (£[redacted] per QALY gained) • Optimistic scenario (£[redacted] per QALY gained) • Pessimistic scenario (£[redacted] per QALY gained)	Report has been amended accordingly.
Section 2.5, Page 38	CIC not marked	Both face-to-face and digital CBT-I (e.g., Sleepio®) have high refusal and failure rates. Among patients who are eligible for CBT-I, only [redacted] achieve the desired results	Report has been amended accordingly.
Section 3.1.2, Page 43	CIC marked incorrectly	The EAG reviewed the clinical study reports (CSRs) of studies 301 and 302 and were surprised to see that the vast majority (87.9%) of the trial populations had not been offered or were not aware of CBT.	Report has been amended accordingly.
Section 3.2.5.1.1, Page 68	CIC not marked	• VAS quality of sleep: MD (95% CI): [redacted] • VAS depth of sleep: MD (95% CI): [redacted] • VAS daytime alertness: MD (95% CI): [redacted] • VAS ability to function: MD (95% CI): [redacted] • PGA-S (daytime symptoms): MD (95% CI): [redacted] • PGA-S (night-time symptoms): [redacted] • PGI-C (daytime symptoms): MD (95% CI): [redacted] • PGI-C (night-time symptoms): MD (95% CI): [redacted] • Mean number of PSG awakenings over night: MD (95% CI): [redacted] • Mean number of self-reported awakenings: MD (95% CI): [redacted]	Report has been amended accordingly.

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
Section 3.2.5.1.2, Page 72	CIC not marked	• PGA-S daytime: MD (95% CI): [redacted] • PGI-C daytime: MD (95% CI): [redacted]	Report has been amended accordingly.
Section 3.2.5.1.2, Page 72	CIC not marked	It can be seen that for PGI-C daytime, the 95% confidence intervals of the small MD of [redacted] cross the null line, indicating a probability of >0.05 that the population MD may not be in the same direction of effect as the point estimate.	Report has been amended accordingly.
Section 3.2.5.1.2, Page 72	CIC not marked	For the results for symptom improvement in terms of quality of sleep, depth of sleep, daytime alertness, and ability to function, no numerical data are given, and the only information is provided in figures (**Error! Reference source not found.**).	Unclear what needs to be amended. Figure 3.1 has been highlighted with a yellow outline, which indicates it is CIC marked.
Section 3.2.5.2.1, Page 74	CIC marked, not underlined	Between-arm analyses were not conducted by the company, so the EAG has carried these out below. The MD (95% CIs) for the two outcomes are as follows: • Change from baseline to Month 3 in latency LPS to REM [MD (95%)]: [redacted]	Report has been amended accordingly.

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
		• Latency sleep onset to REM [MD (95%)] at Month 3: [redacted]	
Section 3.2.5.2.1, Page 74	CIC marked, not underlined	A between-arm analysis was not conducted by the company, so the EAG has carried this out as follows: the MD (95% CIs) is [redacted]	Report has been amended accordingly.
Section 3.2.5.2.1, Page 74	CIC not marked	Numerically, [redacted] from baseline in sleep onset latency (duration from lights off to the first epoch (i.e., 30 seconds) of sleep stage 2 (S2), slow wave sleep (SWS), or REM, or the first 3 consecutive epochs (i.e., 1.5 minutes) of sleep stage 1 (S1)) were observed in participants on daridorexant 50 mg than on placebo, however no statistical comparisons were done (**Error! Reference source not found.**).	Report has been amended accordingly.
Section 3.2.5.2.1, Page 75	CIC marked, not underlined	As a between-arm analysis was not conducted by the company, the EAG has carried this out as follows: the mean difference (MD) (95% CIs) is [redacted]	Report has been amended accordingly.
Section 3.2.5.3.1, Page 75	CIC marked, not underlined	Subjective latency to sleep onset (sLSO) is regarded as a measure of a change in quality of sleep. Daridorexant was observed to lead to [redacted] in sLSO after 3 months compared to placebo (**Error! Reference source not found.**).	Report has been amended accordingly.
Section 3.2.5.3.2, Page 76	CIC marked, not underlined	At 12 months, the difference in improvement between the arms was [redacted] (**Error! Reference source not found.**).	Report has been amended accordingly.

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
Section 3.2.5.4.2, Page 76	CIC marked, not underlined	At 12 months, the [redacted] for the daridorexant arm over placebo for the total score and each of the three domains of the ISDIQ persisted (Table Error**! No text of specified style in document.**.1).	Report has been amended accordingly (p77),

Section 3.2.5.4.2, Page 78	CIC marked, not underlined	Table **Error! No text of specified style in document.**.1: IDSIQ sleepiness domain score, IDSIQ total score, IDSIQ alert/cognition domain score, and IDSIQ mood domain score from baseline to month 12	Report has been amended accordingly.

Visit / **Treatment group**	**n**	**LSM (95% CL)**	**Difference to placebo**	
			LSM (95% CL)	**p-value (two-sided)**
Between treatment analysis for change from baseline in IDSIQ sleepiness domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ total score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ alert/cognition domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]
Between treatment analysis for change from baseline in IDSIQ mood domain score to Month 12				
Daridorexant 50 mg (N=137)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo (N=128)	[redacted]	[redacted]	[redacted]	[redacted]

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
Section 3.2.5.5.1, Page 79	CIC not marked	• WASO: [redacted] • LPS: [redacted] • sTST: [redacted]	Report has been amended accordingly.
Section 3.2.5.5.2, Page 80	CIC not marked	• sTST:[redacted] • This confirms [redacted] between arms	Report has been amended accordingly.
Section 3.2.5.6.2, Page 81	CIC marked, not underlined	There was a [redacted] in ISI© score in the daridorexant arm over the 12 months of Study 303[18] (**Error! Reference source not found.**).	Report has been amended accordingly.
Section 3.2.5.6.2, Page 81	CIC marked, not underlined	No between-arm analysis was carried out by the company. A between-arm analysis carried out by the EAG showed that the difference between arms was [redacted]	Report has been amended accordingly.
Section 3.2.5.6.2, Page 81	CIC marked, not underlined	As a decrease in score is an improvement, this represents a [redacted] for daridorexant.	Report has been amended accordingly.
Section 3.2.5.6.2, Page 81	CIC marked, not underlined	A between-arm analysis carried out by the EAG showed that the relative risk (RR) between arms was [redacted]	Report has been amended accordingly.

<table>
<tr><th>Location of incorrect marking</th><th>Description of incorrect marking</th><th>Amended marking</th><th>EAG comments</th></tr>
<tr><td>Section 4.2.6, Page 101</td><td>CIC not marked</td><td>As the company considered dropout rates from Study 303 were [redacted] clinical practice more accurately than the ones from Study 301, the company provided a scenario analysis using “similar” dropout rates in the first and third month (the ones which were based on Study 301).</td><td>Report has been amended accordingly.</td></tr>
<tr><td>Section 4.2.7, Page 101</td><td>CIC not marked</td><td>Table Error! No text of specified style in document..2: AEs reported at least once in either treatment group
<table>
<tr><th rowspan="2">Treatment-emergent adverse event</th><th colspan="2">Study 301</th><th colspan="2">Study 303</th></tr>
<tr><th>Daridorexant 50 mg N=308; n (%)</th><th>Placebo N=309; n (%)</th><th>Daridorexant 50 mg N=137; n (%)</th><th>Placebo N=128; n (%)</th></tr>
<tr><td>Subjects with at least one TEAE</td><td>116 (37.7)</td><td>105 (34.0)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Subjects with at least one SAE</td><td>3 (1.0)</td><td>7 (2.3)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Subjects with at least one AESIs</td><td>2 (0.6)</td><td>1 (0.3)</td><td>[redacted]</td><td>[redacted]</td></tr>
</table></td><td>Report has been amended accordingly.</td></tr>
<tr><td>Section 5.2, Page 109</td><td>CIC not marked</td><td>• Inclusion of indirect costs (£[redacted] per QALY gained)
• Optimistic scenario (£[redacted] per QALY gained)
• Pessimistic scenario (£[redacted] per QALY gained)</td><td>Report has been amended accordingly.</td></tr>
</table>

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comments
Section 6.4, Page 118	CIC not marked	The ICER [redacted] by ~£[redacted] in the scenario assuming alternative dropout rates.	Report has been amended accordingly.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Daridorexant for treating insomnia [ID3774]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **16 January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Idorsia Pharmaceuticals UK Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Issue 1: It is possible that the population in the trials is narrower than the population in the decision problem. This has implications for the applicability.	No	The population specified in the decision problem is adults with insomnia disorder. This is based on the definition provided by Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®), which defines insomnia disorder as *"dissatisfaction with sleep quantity or quality, associated with difficulty initiating or maintaining sleep, or early morning awakening. Furthermore, the sleep disturbance is associated with significant social or functional distress or impairment. Sleep difficulty occurs at least 3 nights per week and is present for at least 3 months, and occurs despite adequate opportunity for sleep"* (Substance Abuse and Mental Health Services Administration, 2016). Further, the DSM-5® criteria of insomnia disorder is largely consistent with the patient population indicated in the Summary of Product Characteristics (SmPC) for daridorexant, and the same DSM-5® criteria has been used to enroll patients in the pivotal trials of daridorexant. The study population included patients considering these diagnostic criteria, including substantial distress or impairment in social, occupational, educational, academic, behavioural, or other important areas of functioning. Thus, this information supports that

		the study population is representative of the clinical definition of insomnia disorder and of the population in the decision problem.
Issue 2: Although daridorexant is designed as a replacement treatment for those people that may be unsuitable for established care treatments such as CBT-I, most of those in the trials have never had the opportunity to receive or reject CBT-I.	Yes	It is true that the utilisation of CBT-I in the studies is low however this reflects the reality of accessibility and availability of CBT-I. The insomnia disorder patient journey in the UK shows that most patients do not have access to CBT-I. In addition to the 1,000 GP survey presented in the CS (medeConnect Healthcare Insight, 2022), we have sought further confirmatory data from a recent analysis of North Central London Clinical Commissioning Group patient population. It shows insomnia prevalence of [redacted] (North London Partners, 2022). In terms of whether prior exposure to CBT-I (or sleep hygiene) would impact the efficacy of daridorexant, it is important to consider the inclusion criteria of study 301 (Idorsia Pharmaceuticals Ltd – Clinical Study Report for Study 301, 2020). These overlap with many of the sleep hygiene/CBT-I concepts, i.e., keeping a sleep diary, low caffeine/alcohol intake, anchoring bedtime to 9.30pm – 12.30am, having time in bed between 6 and 9 hours (Rossman, 2019; Morin, 2015). This ensures that the population is representative of being exposed to behavioural change after sleep hygiene, the key component of CBT-I, and represents a conservative approach to study daridorexant's effect. In the CS, positioning daridorexant after CBT-I has been offered includes not only patients who have completed CBT-I, but also those who do not have access to, are unsuitable for, have failed, or have refused CBT-I. The population of study 301 reflects this population.

		The 301 inclusion criteria and the fact that subjects have had insomnia diagnosed >10 years ago means most of them will have been exposed to sleep hygiene and with that, the main concepts of CBT-I, more than once.
Issue 3: The comparator in the decision problem is established clinical management. However, the comparator in the clinical effectiveness evidence presented in the CS is placebo with no mention of established clinical management and in the cost effectiveness section it is referred to as no treatment. There is no attempt by the company to perform an indirect treatment comparison to rectify this situation. The CS therefore fails to present data relating to the decision problem.	No	The position regarding active comparators should reflect the scoping and DPM decisions as highlighted in the factual accuracy check document and re-stated below. Extract from pages 4 and 5 of the scoping document ***"It was noted by clinical experts and patient group representatives that access to CBT-I varies geographically. They also noted that CBT-I is not always offered. At the workshop, attendees also discussed pharmacological treatments for insomnia. They explained that none of the currently approved pharmacological treatments are recommended for long term use. Daridorexant is expected to be used to treat insomnia disorder, where symptoms last for more than 3 months per clinical trial. Therefore, the attendees agreed that none of the comparators listed in the draft scope are relevant. The scope has been updated to remove the comparators"*** The following comparators were listed in the draft scope: • Established clinical management (including sleep hygiene and CBT-I) • Zolpidem and zopiclone • Melatonin (for those aged 55 and over) • Benzodiazepines (for example nitrazepam, loprazolam, lormetazepam, temazepam)

		The final scope was updated after the scoping meeting, where all comparators were removed. Whilst established clinical management remained, CBT-I was removed: • Established clinical management (including sleep hygiene advice) without daridorexant Given the context above, established clinical management is interpreted in the CS as placebo in the clinical trial programme as participants in both arms were exposed to sleep hygiene measures during the study. **This was presented and agreed in the DPM and re-iterated during the TEM**. **This is a fundamental criticism that the CS does not answer the decision problem and is repeated throughout the EAG report. This is incorrect and the issue regarding comparators should be considered based on previous discussions and agreements in scoping and in the DPM. The company requests that papers to committee reflect the scoping and subsequent discussions and not the erroneous EAG assumption on comparators.** In addition, the company would like to highlight that the long-term use of existing pharmacological therapies is against all current clinical guidance and expert opinion for the treatment of insomnia disorder. The use of these therapies, beyond their recommended licensed duration, is identified as a significant clinical problem with specific medicines management guidance published by NICE in 2019, 'Benzodiazepine / Hypnotics de-prescribing' with further guidance published in 2022, 'Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults' to reduce prescribing and withdraw patients.

		It cannot be considered established clinical management as a comparison, and this was agreed during scoping and in DPM. Notwithstanding this response we have conducted further analysis linked to issue 9 regarding comparators.
Issue 4: Numerous outcomes that measure the same construct are presented, increasing the risk of type I errors	No	Insomnia disorder is a medical condition diagnosed based on the subjective report of patient's dissatisfaction with sleep and daytime functioning impairment. This supports the inclusion of subjective outcomes of sleep quality/quantity, as they represent the real clinical perspective. On the other hand, objective outcome measures are necessary for efficacy and safety purposes of pharmacological trials. Moreover, it is well known in the clinical scientific community that there is a considerable gap between subjective and objective measures in patients suffering from insomnia disorder. As highlighted in the clarification letter (A20), given the complexity of assessing treatment outcomes in this disorder, it is challenging to prioritise all other outcomes within each category of the NICE final scope since all outcomes within a category should be considered in totality and therefore carry equal importance when evaluating the clinical benefit of daridorexant. Moreover, additional analyses carried out by the EAG on the secondary and exploratory endpoints indicate that most outcomes were in favour of daridorexant (Section 3.2.5 of EAG report). The company proposed prioritising ISI© since it is used as the key effectiveness parameter of the cost-effectiveness model.
Issue 5: The clinical effectiveness evidence (albeit evidence that covers daridorexant versus placebo rather than daridorexant versus established clinical management) omits a key paper	No	This was discussed in the TEM and broadly accepted. The study referred to (201) was a short-term dose response study across four doses of daridorexant (5, 10, 25, or 50mg). The outcomes were assessed on days 1 and 2 only and not deemed relevant to long term treatment of insomnia disorder. The 25mg and 50mg dose were selected for further study in the phase 3 trials.

Issue 6: Ethnic make-up of the trials differs from the ethnic make-up of the UK population. The trials have not been sub-grouped for ethnicity sufficiently comprehensively across the two trials, making it difficult to exclude ethnicity as an effect modifier. Therefore, applicability of the trial findings is unclear.	Yes	The following pharmacokinetic studies on daridorexant considered ethnicity as a covariate: 1. Daridorexant is almost completely metabolised by a single enzyme – 90% attributed to CYP3A4. This enzyme is not prone to polymorphisms across ethnicities (Chaudhry, 2008). 2. The abuse liability study showed that Black/African-Americans metabolise daridorexant faster, but only less than 20% and thus not clinically relevant (Idorsia Phamaceuticals Ltd, 2020). In addition, based on the absence of next-day residual effects in studies 301 and 303, the clinical relevance of differences in daridorexant exposure between subjects with different characteristics (including ethnicity) are expected to be negligible.
Issue 7: Shorter term benefits of daridorexant over placebo do not appear to persist into the longer term in all cases	Yes	The CS utilises 12-month trial data in the economic model and therefore any waning effect is not relevant. It is also clear from the washout period between studies 301 and 303, where patients returned to baseline, that the treatment benefit from daridorexant occurs while taking the drug and that treatment effect stops when treatment stops (Kunz, 2022). This is further discussed in Other Issue 2 below.
Issue 8: Studies 301 and 303 which inform the health economic model excluded patients with mental health problems. Because insomnia is frequently comorbid with other mental health problems the exclusion of patients with mental health problems may decrease the generalisability of the underlying evidence to the decision problem	No	The company acknowledges this limitation, but would like to highlight the potential challenges of including patients with severe or unstable mental health problems in studies 301 and 303. Enrolling patients with comorbid mental health problems in need of treatment can pose a challenge when separating the benefits of daridorexant from that of treatments for mental health problems. These medications are known to affect sleep architecture and previously have been associated with insomnia. Furthermore, they modulate neurotransmitters involved in the regulation of the sleep-wake cycle.

		These supports the decision to exclude patients suffering from unstable or severe mental health problems in the context of efficacy and safety pharmacological trials, as acknowledged by the EAG.
Issue 9: A variety of pharmaceuticals and therapies are available for the treatment of insomnia. The company only included no-treatment as a comparator to daridorexant in the health economic model.	Yes	The original CS presented to NICE calculated the cost-effectiveness of daridorexant compared to no treatment. No treatment was considered to be the appropriate comparator based on the fact that daridorexant is licenced for long-term treatment of insomnia (with appropriate clinical review) and no other pharmacotherapies have a long-term licence for this condition. Although psychoactive drugs such as benzodiazepines and Z-drugs are indicated for short-term alleviation of chronic insomnia, NICE's guidance in its Clinical Knowledge Summary (CKS) on the management of long-term insomnia clearly specifies that long-term use is not indicated due to concerns over issues of increasing dependence and their safety profile: "Do not prescribe long-term hypnotic treatment — for information on withdrawal of hypnotic medication, see the CKS topic on benzodiazepine and Z-drug withdrawal." As the above quote shows, the CKS recommends withdrawal of psychoactive treatment for those who are using it long -term against clinical advice. Adverse events associated with psychoactive drugs include over-sedation, cognitive impairment (with potential link to dementias), increased aggression leading to potential for self-harm/harm to others, and increased accidents (falls and road traffic accidents). Furthermore, due to increasing tolerance, evidence suggests that the clinical effectiveness of psychoactive drugs diminishes with time. That psychoactive drugs are not recommended for long-term use in NICE's own guidance led to the removal of benzodiazepines and Z-

		drugs from the initial scope of the appraisal with the agreement that daridorexant would be the first-line long-term pharmacological treatment for chronic insomnia. Despite the removal of psychoactive drugs from the agreed scope, the EAG commented that these drugs are actively used in the NHS, and they should be considered as a comparator in the daridorexant appraisal. This viewpoint was discussed at the TEM with the EAG in December 2022, and it was agreed that the company would submit additional evidence on indirect cost-effectiveness comparison between the modelling in the CS and a recently published report by NICE looking at the cost-effectiveness of CBT-I alongside tapering off for people addicted to the use of benzodiazepines. In the lifetime analysis presented to NICE, the CS estimated that long-term use of daridorexant was associated with [redacted] QALYs at a lifetime additional cost of [redacted] compared to no pharmacological treatment. NICE guideline NG215 estimated that a CBT-I plus tapering off intervention would cost [redacted] per patient but save [redacted] over their lifetime and [redacted] QALYs by [redacted] compared to usual care (i.e., intervention to get people to stop taking benzodiazepines dominates benzodiazepine treatment). However, the overall effectiveness of CBT-I plus tapering off was estimated to be just 34% in NICE's analysis with 66% remaining on benzodiazepine treatment. Scaling these results up to 100% suggests that stopping benzodiazepine treatment altogether is associated with a [redacted] and [redacted] QALYs. Hence, the long-term cost-effectiveness with daridorexant compared to long-term treatment with benzodiazepines can be estimated as costing [redacted] over the lifetime and leads to [redacted] QALYs.

		This would give an incremental cost-effectiveness ratio (ICER) of approximately [redacted] per QALY for daridorexant compared to benzodiazepines in contrast to an ICER of approximately [redacted] compared to no treatment in a lifetime analysis. It should be pointed out that benzodiazepines and Z-drugs are off-patent medications that are relatively cheap. The cost of lifetime treatment with benzodiazepine is estimated to be just £775 from NICE's NG215 economic model – just 20% of the cost-saving associated with withdrawing patients from long-term psychoactive treatment. The other 80% of the cost-saving is related to long-term adverse events avoided. NICE included the cost of long-term cognitive impairment (dementia), hip fracture, fall injuries, and road traffic accidents in its cost-effectiveness model. By comparison, [redacted] Because daridorexant has no evidence of over-sedation during the day and no cognitive impairment issues, not only are the adverse events of psychoactive drugs not relevant to daridorexant, but there is good evidence that productivity is improved. In the daridorexant clinical trial programme, productivity was measured using the Sheehan Disability Scale (SDS) and [redacted] Although there is no evidence presented in NG215 on productivity effects of psychoactive mediation, the over-sedation and impaired cognitive side effects will worsen productivity scores compared to no treatment. [redacted] (as measured by SDS) in the daridorexant cost-effectiveness analysis highlights the [redacted]

		[redacted]
Issue 10: The company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation.	No	MHRA/EMA have agreed 50 mg of daridorexant is the recommended treatment dose. The 25 mg dose is indicated for a sub-group with a pharmacokinetic issue due to liver dysfunction or co-administration of CYP3A4 inhibitors. This is to achieve "50mg equivalent" daridorexant plasma levels. Therefore, the benefits of daridorexant, and consequently its cost-effectiveness, are expected to be the same for both dosages considering the indicated populations for each of them. [redacted] In addition, while studies 301 and 303 randomised patients to receive the 25 mg dosage, these patients are not reflective of those mentioned in the SmPC. Section 5.2 of the SmPC reflects the results of pharmacokinetic studies measuring exposure to daridorexant after a single dose in patients with liver dysfunction or co-administration of CYP3A4 inhibitors. Due to the abovementioned reasons, the company did not include the 25 mg dosage of daridorexant in the cost-effectiveness model. This was discussed and agreed in the TEM with NICE and the EAG.
Issue 11: As per the CS, the no-treatment arm was modelled to have no dropout, as patients receiving could not dropout from receiving no treatment. However, in the economic model provided by the company, the dropout rates observed in studies 301 and 303 for the daridorexant arm were applied to both daridorexant and no-treatment groups. This	No	The EAG believes that the company has incorrectly included drop out from 'no treatment'. This was not an accurate representation of the company's economic model. During the TEM, the EAG was able to demonstrate the application of their alternative assumptions around dropout. In the spreadsheet provided by the EAG it was noted that the first two 'alterations' proposed by the EAG together (scenarios EAG1 & EAG2) reproduced exactly the ICER of the original CS. This equivalence between the proposed changes by the EAG and the CS was

contradicts the statement made by the company.		confirmed in the algebraic solutions submitted to NICE in the company's factual accuracy check. It was agreed between the EAG and the company at the TEM that the original CS was in fact correct. However, it remains the case that in the report to NICE the presentation of the EAG report refers to the issue as an 'error' on the part of the company. Having agreed there is no error at the TEM, the company requests that the EAG removes 'Issue 11' from the report and all reference to an 'error' being made in the CS as the EAG now agrees that this is factually incorrect. For example: Table 1.12 Key Issue 11 repeats the incorrect statement that dropout was applied to no treatment in the CS (it was not) and then goes on to state that the expected effect on cost-effectiveness is that 'Daridorexant is dominated by no treatment' (page 20). This statement is incorrect (how can a treatment be more effective than no treatment when there is no dropout but less effective than no treatment once dropout is adjusted for?). The mistake that the EAG have made is to implement one half (scenario labelled EAG1) of a different solution to adjust for dropout which gives identical results to the CS when both EAG1 and EAG2 are implemented. When only EAG1 is implemented without the adjustment to utility in EAG2, patients who dropout from treatment get a utility score of 0 which leads to the (erroneous) conclusion that daridorexant is dominated. Section 4.2.6 (p102) continues to misrepresent the situation by stating:

		"*However, in the economic model provided by the company, dropout rates are applied to the incremental values (i.e., the difference between the daridorexant and no-treatment group), instead of being applied to the daridorexant group alone. The EAG will explore applying only the dropout rates to the daridorexant group in their base case"* In Section 6 reporting the EAG's additional analysis, the EAG refers to section 4.2.6 and states (p112) *"Although the company stated in the CS that patients could not dropout from no-treatment, this was not implemented in the economic model as such and hence corrected by the EAG"* in a section labelled '*Fixing errors'* which they describe as *"correcting the model where the company's submitted model was unequivocally wrong"* Table 6.1 includes the statement: *"In contrast to what was stated in the CS, the no-treatment arm dropout rates observed in Study 301 and Study 303 were also applied to the no-treatment arm in the economic model."* This statement is factually incorrect – no dropout rates were applied to the no-treatment arm. Table 6.2 reports the scenario EAG1 alone under the heading 'Fixing errors' resulting in the statement that 'daridorexant is dominated by no treatment'. As argued above this scenario is illogical and involves assigning a zero-utility score to patients who dropout from treatment.

		Having an 'error' flagged by the EAG undermines confidence in the company's model and is prejudicial to the committee. As demonstrated conclusively at the TEM, no such error exists and all reference to this error should be removed from the report.
Issue 12: For the company base case placebo effect was only included for the first three months in the no-treatment arm, but not for the remaining 40 weeks. The EAG considers that the effect of selective attrition on the daridorexant group and the possibility of regression to the mean on the no-treatment group, were not sufficiently justified by the company, and these effects could have biased the comparison in favour of the intervention.	Yes	The company disagrees with the EAG's view that the effect of selective attrition on the daridorexant group and the possibility of regression to the mean on the no-treatment group could have biased the comparison in favour of daridorexant. **The EAG's statement that there was no placebo correction after three months is factually incorrect.** As elaborated in the factual accuracy check (issue 3), in the base case model (Fig 4.1 of EAR), placebo effect was applied to the no-treatment group using study 301 and the same rate was extrapolated to the entire duration of the model. As explained in Document B (B.3.3.2), it was assumed that the no-treatment group would continue at the same ISI© achieved by the end of study 301 (i.e., the 3rd month). This approach is supported by data from study 303 where rebound of ISI© was observed during placebo run out in study 301 prior to study 303 entry. For subjects who continued from study 301 to study 303, a similar placebo effect was subsequently observed in the first 3 months of study 303 (Figure 13, Document B, B.2.9.1). Therefore, it can be considered that study 303 is placebo corrected up to the first 3 months and selective attrition was applied thereafter in the economic model. Selective attrition was proposed as the explanation for the improving ISI© score in study 303 beyond the first three months. The EAG countered that this could be explained by regression to the mean.

		Having set out the clarification on placebo correction, for regression to the mean to be explaining the selective attrition in study 303 after the four-week run in before randomisation into 301 (where some modest regression to the mean was observed), followed by 12 weeks in study 301 and a further 14 weeks in study 303 is highly implausible. The design of studies 301 and 303 largely eliminates any likelihood of regression to the mean. Selective attrition, as implemented in the model is a reasonable and likely explanation for the phenomenon whereby persistence with a daily treatment will be concentrated in those receiving the most benefit from treatment. In addition a recent meta-analysis has characterised the dynamic placebo effect in chronic insomnia, for both within-treatment and after-treatment windows (Jiang, 2020). It clarifies that the placebo effect reaches a stable plateau after 9-12 weeks and is broadly supportive of the approach taken by the company.
Issue 13: The company excluded the AEs reported in studies 301 and 303 from their cost effectiveness model, assuming that these are minor AEs and would not be expected to have consequences on resource use or HRQoL.	No	When the EAG originally raised this issue, the company responded with a simplified illustration to show that minor AEs would not change the ICER substantially. This is a form of a fortiori analysis that makes the argument 'even if' the AE had a large impact on HRQoL then it is unlikely to impact the estimated cost-effectiveness because the differences in AEs between placebo and treatment groups in the trial were so small (as acknowledged in the label for daridorexant). While the company accepts the desire of the EAG to see accurate modelling of the AEs – the company trusts that the a fortiori analysis is nevertheless sufficient to convince NICE and the committee that AEs are not an important part of the cost-effectiveness story. If a further rationale is needed then it would be that with a once daily treatment, if side effects were to be anything other than minor then it is likely patients would discontinue treatment.

Issue 14: There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.	Yes	The original response to the CS by the EAG included some very helpful suggestions for improving the ISI©-EQ5D mapping. The company has taken the opportunity to incorporate these suggestions and has produced a manuscript detailing the results which is accepted for publication at *PharmacoEconomics-Open*. As suggested by the EAG, additional modelling types were tested, and an additional validation exercise was undertaken. Although the suggested CLAD model performed poorly, the suggested ALDVMM model performed well – outperforming the original gamma-log GLM by a small margin. The submitted manuscript is appended as additional information (Chalet, 2023) and the CEM is now updated to include ALDVMM as a scenario which increases the base case ICER by a small amount.
Issue 15: In addition to treatment acquisition costs, the CS only incorporated costs and resource use for GP visits, emergency room visits and inpatient care. The company justified the decision due to these being the only categories captured in the NHWS dataset and stating that the approach was conservative. Such a conclusion cannot be drawn in the absence of supporting evidence.	No	The company disagrees with the EAG over their interpretation. As stated, the NHWS dataset only included GP visits, emergency room visits, and inpatient hospital stays. For all three categories there was a positive association between higher ISI© score (more severe insomnia) and higher resource use. The main categories of resource use missing from NHWS that would usually be included in a health service perspective/NICE Reference Case analysis are concomitant medications and outpatient attendances. Even without evidence, it is unlikely that these excluded categories would have an inverse relationship with ISI©. Assuming a positive association with ISI© implies the CS is conservative by excluding these categories of cost since reduction in insomnia with effective treatment would lead to further savings in these cost categories.
Other issues identified by NICE technical team (not included in the EAR):		
Other Issue 1: Time horizon not adequately long enough to capture long-term costs and outcomes. Company suggest that daridorexant	No	In the decision problem meeting, the submission and in the TEM, the company repeated its assertion that a one-year model is adequate to

will only be provided for a maximum of 1 year, but no formal stopping rule exists. A potential risk that the treatment is provided for longer than 1 year and has not been appropriately modelled.		estimate cost-effectiveness. There are several reasons for this that have been argued consistently by the company: • Pharmacodynamics of treatment such that the benefits of treatment apply and are lost within hours of taking/ stopping treatment. • On this basis, a single day would be sufficient to estimate the impact of a short acting treatment that is taken once-a-day. However, a longer period allows for dropout adjustment – particularly allowing patients for whom treatment is not working so well to be part of the natural attrition. • A one-year timeframe corresponds to the combined period of the studies 301 and 303, representing the best evidence for any insomnia treatment in the medium term. This allows important parameters (including dropout) to be observed as well as to demonstrate that treatment effect is maintained into the medium term. • A one-year timeframe is conservative from the perspective of estimated cost-effectiveness because it includes wasted prescriptions for those who do not take their medicine and that there is heterogeneity in treatment response. The desire for a longer-term model implied by the wording of this issue is simply that treatment may last for longer than one year. Despite the statement that this has not been appropriately modelled, the CS did include a longer-term model as a scenario rather than a base case. This is because, as demonstrated in the 301 and 303 studies, patients persisting with treatment to one year have better outcomes on average and based on the epidemiological relationship between poor sleep and poor long term health outcomes (increased risk of cardiovascular disease), the cost-effectiveness of long-term treatment improves on the one-year base case estimates presented.

Other Issue 2: Extrapolation of short-term benefits and any potential waning of treatment effect over the long term (relating to Other Issue 1 and EAG Issue 7).	No	As described in relation to Other Issue 1, a lifetime model of long-term daridorexant lowers the estimated ICER due to selective attrition and potential long term health benefits. There was no waning in the treatment effect on ISI© observed over 12 months and no reason (such as increased tolerance as with the benzodiazipenes and Z drugs) to expect this in the future. Therefore, there is no basis to estimate waning of treatment effect. In any case the importance of treatment waning in the future is most relevant to treatments which have stopped (for example – do treatment effects of a new cancer drug continue beyond the trial after treatment is discontinued?). For a once-a-day treatment like daridorexant, where the pharmacodynamics and pharmacokinetics support early onset of effect and early washout, any treatment waning should be noticed by the patient and should lead to treatment discontinuation – instigated either through the patient discontinuing themselves or by a treating physician at one of the recommended treatment review visits.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response

Additional issue 1: The EAG report notes uncertainty in the approach to exploring the impact of productivity losses as we explore two methods with different results, but does not sufficiently highlight the one directional positive impact and associated societal value of treating insomnia disorder.	Section 4.2.9.5	Yes	Whilst not raised as a specific issue we have included scenario analyses on indirect costs per "NICE health technology evaluations: the manual". Chronic insomnia impacts an individual's mental and physical health, quality of life (QoL), and productivity. Importantly, the consequences of insomnia go well beyond the individual, as there may be cascading effects on families, employers, and global economies. This is increasingly recognised, and the productivity impact is therefore a critical component of the overall cost-effectiveness of interventions for chronic insomnia. **There are certainly arguments that this should be the base case in our submission however we hope to make the case that the scenario with productivity included can be presented to the committee per section 4.4.23 of the NICE Manual on productivity costs - "They can be presented separately, as additional information for the committee, if such costs may be a critical component of the value of the technology."** When these costs are included directly using the SDS measured from the clinical trial, treatment with daridorexant is effectively cost neutral in year 1 and cost saving in subsequent years. When they are estimated indirectly from WPAI mapped to ISI©, daridorexant is cost saving after 3 months of treatment. In our submission we reference RAND Europe "Why Sleep Matters: Quantifying the Economic Costs of Insufficient Sleep" There is now an additional report from RAND Europe [redacted]

			[redacted] (RAND, 2023). We have permission to share the Executive Summary as a confidential document and this is attached as a reference. This report highlights [redacted] This supports the broader case on the economic value of treating chronic insomnia.

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Issue 14: There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.	In the CS, the base case ICER estimated using the original ISI©-EQ5D mapping and gamma-log GLM was [redacted]/ QALY.	As per the response to Issue 14, the company has incorporated the EAG's suggestions to improve the ISI©-EQ5D mapping. Additional modelling types were tested (CLAD and ALDVMM) and an additional validation exercise was undertaken. Although the suggested CLAD model performed poorly, the suggested ALDVMM model performed well – outperforming the original gamma-log GLM by a small margin. The CEM is now updated to include ALDVMM as a scenario which increases the base case ICER by a small amount.	Base-case ICER resulting from the change: [redacted]/ QALY Change from CS base-case ICER: [redacted]/ QALY

Sensitivity analyses around revised base case

PLEASE DESCRIBE HERE

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Daridorexant for treating insomnia [ID3774]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **16 January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Idorsia Pharmaceuticals UK Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response	EAG comment
Issue 1: It is possible that the population in the trials is narrower than the population in the decision problem. This has implications for the applicability.	No	The population specified in the decision problem is adults with insomnia disorder. This is based on the definition provided by Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5®), which defines insomnia disorder as *"dissatisfaction with sleep quantity or quality, associated with difficulty initiating or maintaining sleep, or early morning awakening. Furthermore, the sleep disturbance is associated with significant social or functional distress or impairment. Sleep difficulty occurs at least 3 nights per week and is present for at least 3 months, and occurs despite adequate opportunity for sleep"* (Substance Abuse and Mental Health Services Administration, 2016).	Uncertainty persists because there remain some inclusion criteria for the trial that are not covered by the characteristics provided by the company that define 'adults with insomnia disorder'. For example, the trial participants are limited to those that take >30 minutes to fall asleep, but it is unclear that taking ≤30 minutes to fall asleep would define someone as not having insomnia disorder. Given the possibility that the population in the trials is narrower than the population in the decision problem, there are implications that the trial results might not necessarily be applicable to the clinical population. This therefore remains as a key issue.

		Further, the DSM-5® criteria of insomnia disorder is largely consistent with the patient population indicated in the Summary of Product Characteristics (SmPC) for daridorexant, and the same DSM-5® criteria has been used to enrol patients in the pivotal trials of daridorexant. The study population included patients considering these diagnostic criteria, including substantial distress or impairment in social, occupational, educational, academic, behavioural, or other important areas of functioning. Thus, this information supports that the study population is representative of the clinical definition of insomnia disorder and of the population in the decision problem.	
Issue 2: Although daridorexant is designed as a replacement treatment for those people that may be unsuitable for established care treatments such as CBT-I, most of those in the trials have never had the opportunity to receive or reject CBT-I.	Yes	It is true that the utilisation of CBT-I in the studies is low however this reflects the reality of accessibility and availability of CBT-I. The insomnia disorder patient journey in the UK shows that most patients do not have access to CBT-I. In addition to the 1,000 GP survey presented in the CS (medeConnect Healthcare Insight, 2022), we have sought further confirmatory data from a recent analysis of North Central London Clinical Commissioning Group patient population. It shows insomnia disorder? prevalence of [redacted] (North London Partners, 2022).	The EAG is grateful for the additional information that the company gained from the London Clinical Commissioning group. This does back up the company's statement that the low level of CBT-I usage in the trial is consistent with that of the UK target population. This is therefore no longer a key issue.

		In terms of whether prior exposure to CBT-I (or sleep hygiene) would impact the efficacy of daridorexant, it is important to consider the inclusion criteria of study 301 (Idorsia Pharmaceuticals Ltd – Clinical Study Report for Study 301, 2020). These overlap with many of the sleep hygiene/CBT-I concepts, i.e., keeping a sleep diary, low caffeine/alcohol intake, anchoring bedtime to 9.30pm – 12.30am, having time in bed between 6 and 9 hours (Rossman, 2019; Morin, 2015). This ensures that the population is representative of being exposed to behavioural change after sleep hygiene, the key component of CBT-I, and represents a conservative approach to study daridorexant's effect. In the CS, positioning daridorexant after CBT-I has been offered includes not only patients who have completed CBT-I, but also those who do not have access to, are unsuitable for, have failed, or have refused CBT-I. The population of study 301 reflects this population. The 301 inclusion criteria and the fact that subjects have had insomnia diagnosed >10 years ago means most of them will have been exposed to sleep hygiene and	

		with that, the main concepts of CBT-I, more than once.	
Issue 3: The comparator in the decision problem is established clinical management. However, the comparator in the clinical effectiveness evidence presented in the CS is placebo with no mention of established clinical management and in the cost effectiveness section it is referred to as no treatment. There is no attempt by the company to perform an indirect treatment comparison to rectify this situation. The CS therefore fails to present data relating to the decision problem.	No	The position regarding active comparators should reflect the scoping and DPM decisions as highlighted in the factual accuracy check document and re-stated below. Extract from pages 4 and 5 of the scoping document ***"It was noted by clinical experts and patient group representatives that access to CBT-I varies geographically. They also noted that CBT-I is not always offered. At the workshop, attendees also discussed pharmacological treatments for insomnia. They explained that none of the currently approved pharmacological treatments are recommended for long term use. Daridorexant is expected to be used to treat insomnia disorder, where symptoms last for more than 3 months per clinical trial. Therefore, the attendees agreed that none of the comparators listed in the draft scope are relevant. The scope has been updated to remove the comparators"*** The following comparators were listed in the draft scope:	The EAG would reiterate that the NICE scope and the decision problem define the proposed comparison as daridorexant versus established clinical management (ECM). This is not the same as daridorexant + ECM versus placebo + ECM, which is what was performed in the trials. This is true regardless of discussions for which the EAG were not involved. However, if the company can show that in clinical practice daridorexant would always be given with other ECM, then it would make sense for the decision problem to have originally been defined as daridorexant + ECM versus placebo + ECM. This would then tally with the trial data. If the committee agrees that the decision problem intervention should be daridorexant + ECM rather than daridorexant, then this should resolve the issue.

		• Established clinical management (including sleep hygiene and CBT-I) • Zolpidem and zopiclone • Melatonin (for those aged 55 and over) • Benzodiazepines (for example nitrazepam, loprazolam, lormetazepam, temazepam) The final scope was updated after the scoping meeting, where all comparators were removed. Whilst established clinical management remained, CBT-I was removed: • Established clinical management (including sleep hygiene advice) without daridorexant Given the context above, established clinical management is interpreted in the CS as placebo in the clinical trial programme as participants in both arms were exposed to sleep hygiene measures during the study. **This was presented and agreed in the DPM and re-iterated during the TEM**. **This is a fundamental criticism that the CS does not answer the decision problem and is repeated throughout the EAG report. This is incorrect and the issue regarding comparators should be**	

		considered based on previous discussions and agreements in scoping and in the DPM. The company requests that papers to committee reflect the scoping and subsequent discussions and not the erroneous EAG assumption on comparators. In addition, the company would like to highlight that the long-term use of existing pharmacological therapies is against all current clinical guidance and expert opinion for the treatment of insomnia disorder. The use of these therapies, beyond their recommended licensed duration, is identified as a significant clinical problem with specific medicines management guidance published by NICE in 2019, 'Benzodiazepine / Hypnotics de-prescribing' with further guidance published in 2022, 'Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults' to reduce prescribing and withdraw patients. It cannot be considered established clinical management as a comparison, and this was agreed during scoping and in DPM. Notwithstanding this response we have conducted further analysis linked to issue 9 regarding comparators.	

Issue 4: Numerous outcomes that measure the same construct are presented, increasing the risk of type I errors	No	Insomnia disorder is a medical condition diagnosed based on the subjective report of patient's dissatisfaction with sleep and daytime functioning impairment. This supports the inclusion of subjective outcomes of sleep quality/quantity, as they represent the real clinical perspective. On the other hand, objective outcome measures are necessary for efficacy and safety purposes of pharmacological trials. Moreover, it is well known in the clinical scientific community that there is a considerable gap between subjective and objective measures in patients suffering from insomnia disorder. As highlighted in the clarification letter (A20), given the complexity of assessing treatment outcomes in this disorder, it is challenging to prioritise all other outcomes within each category of the NICE final scope since all outcomes within a category should be considered in totality and therefore carry equal importance when evaluating the clinical benefit of daridorexant. Moreover, additional analyses carried out by the EAG on the secondary and exploratory endpoints indicate that most outcomes were in favour of daridorexant (Section 3.2.5 of EAG report).	Based on the arguments provided by the company, the EAG is happy to remove this as a key issue.

		The company proposed prioritising ISI© since it is used as the key effectiveness parameter of the cost-effectiveness model.	
Issue 5: The clinical effectiveness evidence (albeit evidence that covers daridorexant versus placebo rather than daridorexant versus established clinical management) omits a key paper	No	This was discussed in the TEM and broadly accepted. The study referred to (201) was a short-term dose response study across four doses of daridorexant (5, 10, 25, or 50mg). The outcomes were assessed on days 1 and 2 only and not deemed relevant to long term treatment of insomnia disorder. The 25mg and 50mg dose were selected for further study in the phase 3 trials.	No comments
Issue 6: Ethnic make-up of the trials differs from the ethnic make-up of the UK population. The trials have not been sub-grouped for ethnicity sufficiently comprehensively across the two trials, making it difficult to exclude ethnicity as an effect modifier. Therefore, applicability of the trial findings is unclear.	Yes	The following pharmacokinetic studies on daridorexant considered ethnicity as a covariate: 1. Daridorexant is almost completely metabolised by a single enzyme – 90% attributed to CYP3A4. This enzyme is not prone to polymorphisms across ethnicities (Chaudhry, 2008). 2. The abuse liability study showed that Black/African-Americans metabolise daridorexant faster, but only less than 20% and thus not clinically relevant (Idorsia Phamaceuticals Ltd, 2020). In addition, based on the absence of next-day residual effects in studies 301 and 303, the clinical relevance of differences in daridorexant exposure between subjects	All of the three examples cited by the company that aim to show that ethnicity is not an effect modifier focus on the effect of ethnicity on the rate of metabolism of daridorexant. However, the differing efficacy or safety of daridorexant across ethnicities likely depends on more factors than just the rate of daridorexant metabolism. This is therefore still a key issue.

		with different characteristics (including ethnicity) are expected to be negligible.	
Issue 7: Shorter term benefits of daridorexant over placebo do not appear to persist into the longer term in all cases	Yes	The CS utilises 12-month trial data in the economic model and therefore any waning effect is not relevant. It is also clear from the washout period between studies 301 and 303, where patients returned to baseline, that the treatment benefit from daridorexant occurs while taking the drug and that treatment effect stops when treatment stops (Kunz, 2022). This is further discussed in Other Issue 2 below.	The EAG agrees with the company's observation that the treatment benefit ceases after stopping the drug. This is actually the point that the EAG is making: that the drug does not appear to have a long-term effect thus may need to be taken indefinitely. This is an important factor that needs to be emphasised to the committee. This is therefore still a key issue.
Issue 8: Studies 301 and 303 which inform the health economic model excluded patients with mental health problems. Because insomnia is frequently comorbid with other mental health problems the exclusion of patients with mental health problems may decrease the generalisability of the underlying evidence to the decision problem	No	The company acknowledges this limitation, but would like to highlight the potential challenges of including patients with severe or unstable mental health problems in studies 301 and 303. Enrolling patients with comorbid mental health problems in need of treatment can pose a challenge when separating the benefits of daridorexant from that of treatments for mental health problems. These medications are known to affect sleep architecture and previously have been associated with insomnia. Furthermore, they modulate neurotransmitters involved in the regulation of the sleep-wake cycle. These supports the decision to exclude patients suffering from unstable or severe mental health problems in the context of	Thank you for providing this additional information. Highlighting the trade-off between study feasibility and representativeness with regards to the exclusion of patients with mental health problems. The EAG believes the key issue, exclusion of patients with mental health problems and the potential decreased generalisability of the underlying evidence to the decision problem, is still relevant for the committee to consider.

		efficacy and safety pharmacological trials, as acknowledged by the EAG.	
Issue 9: A variety of pharmaceuticals and therapies are available for the treatment of insomnia. The company only included no-treatment as a comparator to daridorexant in the health economic model.	Yes	The original CS presented to NICE calculated the cost-effectiveness of daridorexant compared to no treatment. No treatment was considered to be the appropriate comparator based on the fact that daridorexant is licenced for long-term treatment of insomnia (with appropriate clinical review) and no other pharmacotherapies have a long-term licence for this condition. Although psychoactive drugs such as benzodiazepines and Z-drugs are indicated for short-term alleviation of chronic insomnia, NICE's guidance in its Clinical Knowledge Summary (CKS) on the management of long-term insomnia clearly specifies that long-term use is not indicated due to concerns over issues of increasing dependence and their safety profile: "Do not prescribe long-term hypnotic treatment — for information on withdrawal of hypnotic medication, see the CKS topic on benzodiazepine and Z-drug withdrawal." As the above quote shows, the CKS recommends withdrawal of psychoactive	Thank you for providing this additional information. The EAG would prefer that the company's base-case analyses (i.e. with 12 month time horizon) would include all relevant comparators.

		treatment for those who are using it long - term against clinical advice. Adverse events associated with psychoactive drugs include over-sedation, cognitive impairment (with potential link to dementias), increased aggression leading to potential for self-harm/harm to others, and increased accidents (falls and road traffic accidents). Furthermore, due to increasing tolerance, evidence suggests that the clinical effectiveness of psychoactive drugs diminishes with time. That psychoactive drugs are not recommended for long-term use in NICE's own guidance led to the removal of benzodiazepines and Z-drugs from the initial scope of the appraisal with the agreement that daridorexant would be the first-line long-term pharmacological treatment for chronic insomnia. Despite the removal of psychoactive drugs from the agreed scope, the EAG commented that these drugs are actively used in the NHS, and they should be considered as a comparator in the daridorexant appraisal. This viewpoint was discussed at the TEM with the EAG in December 2022, and it was agreed that the company would submit additional evidence on indirect cost-effectiveness	

		comparison between the modelling in the CS and a recently published report by NICE looking at the cost-effectiveness of CBT-I alongside tapering off for people addicted to the use of benzodiazepines. In the lifetime analysis presented to NICE, the CS estimated that long-term use of daridorexant was associated with [redacted] QALYs at a lifetime additional cost of [redacted] compared to no pharmacological treatment. NICE guideline NG215 estimated that a CBT-I plus tapering off intervention would cost [redacted] per patient but save [redacted] over their lifetime and [redacted] QALYs by [redacted] compared to usual care (i.e., intervention to get people to stop taking benzodiazepines dominates benzodiazepine treatment). However, the overall effectiveness of CBT-I plus tapering off was estimated to be just 34% in NICE's analysis with 66% remaining on benzodiazepine treatment. Scaling these results up to 100% suggests that stopping benzodiazepine treatment altogether is associated with a [redacted] QALYs. Hence, the long-term cost-effectiveness with daridorexant compared to long-term treatment with benzodiazepines can be estimated as costing [redacted] over the lifetime and leads to [redacted] QALYs. This would give an incremental cost-	

		effectiveness ratio (ICER) of approximately [redacted] per QALY for daridorexant compared to benzodiazepines in contrast to an ICER of approximately [redacted] compared to no treatment in a lifetime analysis. It should be pointed out that benzodiazepines and Z-drugs are off-patent medications that are relatively cheap. The cost of lifetime treatment with benzodiazepine is estimated to be just £775 from NICE's NG215 economic model – just 20% of the cost-saving associated with withdrawing patients from long-term psychoactive treatment. The other 80% of the cost-saving is related to long-term adverse events avoided. NICE included the cost of long-term cognitive impairment (dementia), hip fracture, fall injuries, and road traffic accidents in its cost-effectiveness model. By comparison, [redacted] Because daridorexant has no evidence of over-sedation during the day and no cognitive impairment issues, not only are the adverse events of psychoactive drugs not relevant to daridorexant, but there is good evidence that productivity is improved. In the daridorexant clinical trial programme, productivity was measured	

		using the Sheehan Disability Scale (SDS) and [redacted] Although there is no evidence presented in NG215 on productivity effects of psychoactive mediation, the over-sedation and impaired cognitive side effects will worsen productivity scores compared to no treatment. [redacted] (as measured by SDS) in the daridorexant cost-effectiveness analysis highlights the [redacted]	
Issue 10: The company did not include the 25 mg dosage of daridorexant in the cost effectiveness model even though it is part of the anticipated market authorisation.	No	MHRA/EMA have agreed 50 mg of daridorexant is the recommended treatment dose. The 25 mg dose is indicated for a sub-group with a pharmacokinetic issue due to liver dysfunction or co-administration of CYP3A4 inhibitors. This is to achieve "50mg equivalent" daridorexant plasma levels. Therefore, the benefits of daridorexant, and consequently its cost-effectiveness, are expected to be the same for both dosages considering the indicated populations for each of them. [redacted] In addition, while studies 301 and 303 randomised patients to receive the 25 mg dosage, these patients are not reflective of those mentioned in the SmPC. Section 5.2 of the SmPC reflects the results of	Although the company provide helpful explanations with regards to the use of the 25 mg dose (related to pharmacokinetic issue). The EAG does not recall that it was agreed not to consider the 25 mg dosage in the cost-effectiveness analyses during the TEM. Given the 25 mg dose is potentially used in clinical practice (where there is co-administration of moderate CYP3A4 inhibitors), providing a scenario analysis using only data from patients who received the 25 mg population in studies 301 and 303 would be of interest (as highlighted in this key issue).

		pharmacokinetic studies measuring exposure to daridorexant after a single dose in patients with liver dysfunction or co-administration of CYP3A4 inhibitors. Due to the abovementioned reasons, the company did not include the 25 mg dosage of daridorexant in the cost-effectiveness model. This was discussed and agreed in the TEM with NICE and the EAG.	
Issue 11: As per the CS, the no-treatment arm was modelled to have no dropout, as patients receiving could not dropout from receiving no treatment. However, in the economic model provided by the company, the dropout rates observed in studies 301 and 303 for the daridorexant arm were applied to both daridorexant and no-treatment groups. This contradicts the statement made by the company.	No	The EAG believes that the company has incorrectly included drop out from 'no treatment'. This was not an accurate representation of the company's economic model. During the TEM, the EAG was able to demonstrate the application of their alternative assumptions around dropout. In the spreadsheet provided by the EAG it was noted that the first two 'alterations' proposed by the EAG together (scenarios EAG1 & EAG2) reproduced exactly the ICER of the original CS. This equivalence between the proposed changes by the EAG and the CS was confirmed in the algebraic solutions submitted to NICE in the company's factual accuracy check. It was agreed between the EAG and the company at the TEM that the original CS was in fact correct. However, it remains the case that in the report to NICE the	The EAG agreed with the company that the first two 'alterations' proposed by the EAG together (scenarios EAG1 & EAG2) reproduced exactly the ICER of the original CS. The description of the fixing error (section 6.1.1.1 of the EAG report) is not factual incorrect. It addressed an inconsistency between the description of the analyses in the CS and the implementation of the economic model, which was considered an error by the EAG. As stated in the EAG report: *"According to the company, treatment discontinuation (i.e., dropout rates) were not incorporated in the no-treatment group as patients receiving no treatment would not be able of dropout from the lack of treatment. However, in the economic model provided by the company, dropout rates are applied to the incremental values (i.e., the difference between the daridorexant and no-*

		presentation of the EAG report refers to the issue as an 'error' on the part of the company. Having agreed there is no error at the TEM, the company requests that the EAG removes 'Issue 11' from the report and all reference to an 'error' being made in the CS as the EAG now agrees that this is factually incorrect. For example: Table 1.12 Key Issue 11 repeats the incorrect statement that dropout was applied to no treatment in the CS (it was not) and then goes on to state that the expected effect on cost-effectiveness is that 'Daridorexant is dominated by no treatment' (page 20). This statement is incorrect (how can a treatment be more effective than no treatment when there is no dropout but less effective than no treatment once dropout is adjusted for?). The mistake that the EAG have made is to implement one half (scenario labelled EAG1) of a different solution to adjust for dropout which gives identical results to the CS when both EAG1 and EAG2 are implemented. When only EAG1 is implemented without the adjustment to utility in EAG2, patients who dropout from treatment get a utility score of 0 which	*treatment group), instead of being applied to the daridorexant group alone."* Without any explanation provided in the CS, the EAG could not 'simply' assume that the difference between the CS and the model implementation was due to a simplification of arithmetic equations. Transparency is a key aspect of modelling and arguably even more important in uncommon analyses approaches such as the 'mediated' analysis adopted by the company. Lack of transparency undermines the EAG's confidence in the company's model. Nevertheless, as stated above, this issue is not influential in terms of impact on the estimated results.

		leads to the (erroneous) conclusion that daridorexant is dominated. Section 4.2.6 (p102) continues to misrepresent the situation by stating: "*However, in the economic model provided by the company, dropout rates are applied to the incremental values (i.e., the difference between the daridorexant and no-treatment group), instead of being applied to the daridorexant group alone. The EAG will explore applying only the dropout rates to the daridorexant group in their base case"* In Section 6 reporting the EAG's additional analysis, the EAG refers to section 4.2.6 and states (p112) *"Although the company stated in the CS that patients could not dropout from no-treatment, this was not implemented in the economic model as such and hence corrected by the EAG"* in a section labelled '*Fixing errors'* which they describe as	

		"correcting the model where the company's submitted model was unequivocally wrong" Table 6.1 includes the statement: *"In contrast to what was stated in the CS, the no-treatment arm dropout rates observed in Study 301 and Study 303 were also applied to the no-treatment arm in the economic model."* This statement is factually incorrect – no dropout rates were applied to the no-treatment arm. Table 6.2 reports the scenario EAG1 alone under the heading 'Fixing errors' resulting in the statement that 'daridorexant is dominated by no treatment'. As argued above this scenario is illogical and involves assigning a zero-utility score to patients who dropout from treatment. **Having an 'error' flagged by the EAG undermines confidence in the company's model and is prejudicial to the committee. As demonstrated conclusively at the TEM, no such error exists and all reference to this error should be removed from the report.**	
Issue 12: For the company base case placebo effect was only included for the	Yes	The company disagrees with the EAG's view that the effect of selective attrition on the daridorexant group and the possibility	It is unclear to the EAG what is exactly factually incorrect according to the company (also given the lack of referral to a specific statement in the

first three months in the no-treatment arm, but not for the remaining 40 weeks. The EAG considers that the effect of selective attrition on the daridorexant group and the possibility of regression to the mean on the no-treatment group, were not sufficiently justified by the company, and these effects could have biased the comparison in favour of the intervention.		of regression to the mean on the no-treatment group could have biased the comparison in favour of daridorexant. **The EAG's statement that there was no placebo correction after three months is factually incorrect.** As elaborated in the factual accuracy check (issue 3), in the base case model (Fig 4.1 of EAR), placebo effect was applied to the no-treatment group using study 301 and the same rate was extrapolated to the entire duration of the model. As explained in Document B (B.3.3.2), it was assumed that the no-treatment group would continue at the same ISI© achieved by the end of study 301 (i.e., the 3rd month). This approach is supported by data from study 303 where rebound of ISI© was observed during placebo run out in study 301 prior to study 303 entry. For subjects who continued from study 301 to study 303, a similar placebo effect was subsequently observed in the first 3 months of study 303 (Figure 13, Document B, B.2.9.1). Therefore, it can be considered that study 303 is placebo corrected up to the first 3 months and selective attrition was applied thereafter in	EAG report). The EAG did not state that there was no placebo correction after three months. The EAG stated that: "The company included the placebo effect reported both in Study 301 and Study 303 only in their pessimistic scenario analysis (this increased the ICER to [redacted] and [redacted] , including the dropout rates). In the company's base case, placebo was only adjusted for the first 3 months in the no-treatment arm, but not thereafter." As stated in the CS (and illustrated in CS Figure 15), the placebo effect observed in Study 303 (after 3 month) was not incorporated in the CS base-case. The EAG had two main issues with the placebo adjustment assumptions and application (as highlighted in section 4.2.6 of the EAG report). Hence, the EAG incorporated placebo adjustment for the analyses time horizon of 12 months (based on both Study 301 and Study 303) in the EAG base case.

		the economic model. Selective attrition was proposed as the explanation for the improving ISI© score in study 303 beyond the first three months. The EAG countered that this could be explained by regression to the mean. Having set out the clarification on placebo correction, for regression to the mean to be explaining the selective attrition in study 303 after the four-week run in before randomisation into 301 (where some modest regression to the mean was observed), followed by 12 weeks in study 301 and a further 14 weeks in study 303 is highly implausible. The design of studies 301 and 303 largely eliminates any likelihood of regression to the mean. Selective attrition, as implemented in the model is a reasonable and likely explanation for the phenomenon whereby persistence with a daily treatment will be concentrated in those receiving the most benefit from treatment. In addition a recent meta-analysis has characterised the dynamic placebo effect in chronic insomnia, for both within-treatment and after-treatment windows (Jiang, 2020). It clarifies that the placebo effect reaches a stable plateau after 9-12	

		weeks and is broadly supportive of the approach taken by the company.	
Issue 13: The company excluded the AEs reported in studies 301 and 303 from their cost effectiveness model, assuming that these are minor AEs and would not be expected to have consequences on resource use or HRQoL.	No	When the EAG originally raised this issue, the company responded with a simplified illustration to show that minor AEs would not change the ICER substantially. This is a form of a fortiori analysis that makes the argument 'even if' the AE had a large impact on HRQoL then it is unlikely to impact the estimated cost-effectiveness because the differences in AEs between placebo and treatment groups in the trial were so small (as acknowledged in the label for daridorexant). While the company accepts the desire of the EAG to see accurate modelling of the AEs – the company trusts that the a fortiori analysis is nevertheless sufficient to convince NICE and the committee that AEs are not an important part of the cost-effectiveness story. If a further rationale is needed then it would be that with a once daily treatment, if side effects were to be anything other than minor then it is likely patients would discontinue treatment.	Thank you for providing further context to the analysis provided in response to clarification question B12 (and highlighted in section 4.2.7 of the EAG report).
Issue 14: There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of)	Yes	The original response to the CS by the EAG included some very helpful suggestions for improving the ISI©-EQ5D mapping. The company has taken the opportunity to incorporate these suggestions and has produced a manuscript detailing the results which is	Thank you for providing this additional information.

conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function and (lack of) exploring other model types.		accepted for publication at *PharmacoEconomics-Open*. As suggested by the EAG, additional modelling types were tested, and an additional validation exercise was undertaken. Although the suggested CLAD model performed poorly, the suggested ALDVMM model performed well – outperforming the original gamma-log GLM by a small margin. The submitted manuscript is appended as additional information (Chalet, 2023) and the CEM is now updated to include ALDVMM as a scenario which increases the base case ICER by a small amount.	
Issue 15: In addition to treatment acquisition costs, the CS only incorporated costs and resource use for GP visits, emergency room visits and inpatient care. The company justified the decision due to these being the only categories captured in the NHWS dataset and stating that the approach was conservative. Such a conclusion cannot be drawn in the absence of supporting evidence.	No	The company disagrees with the EAG over their interpretation. As stated, the NHWS dataset only included GP visits, emergency room visits, and inpatient hospital stays. For all three categories there was a positive association between higher ISI© score (more severe insomnia) and higher resource use. The main categories of resource use missing from NHWS that would usually be included in a health service perspective/NICE Reference Case analysis are concomitant medications and outpatient attendances. Even without evidence, it is unlikely that these excluded categories would have an inverse relationship with ISI©. Assuming a positive association with ISI© implies the CS is conservative by excluding these	Thank you for providing this additional information. Indeed, if there is a positive association between higher ISI© score (more severe insomnia) and higher resource use (for all missing cost items), this can be considered a conservative approach.

		categories of cost since reduction in insomnia with effective treatment would lead to further savings in these cost categories.	
Other issues identified by NICE technical team (not included in the EAR):			
Other Issue 1: Time horizon not adequately long enough to capture long-term costs and outcomes. Company suggest that daridorexant will only be provided for a maximum of 1 year, but no formal stopping rule exists. A potential risk that the treatment is provided for longer than 1 year and has not been appropriately modelled.	No	In the decision problem meeting, the submission and in the TEM, the company repeated its assertion that a one-year model is adequate to estimate cost-effectiveness. There are several reasons for this that have been argued consistently by the company: • Pharmacodynamics of treatment such that the benefits of treatment apply and are lost within hours of taking/ stopping treatment. • On this basis, a single day would be sufficient to estimate the impact of a short acting treatment that is taken once-a-day. However, a longer period allows for dropout adjustment – particularly allowing patients for whom treatment is not working so well to be part of the natural attrition. • A one-year timeframe corresponds to the combined period of the studies 301 and 303, representing the best evidence for any insomnia treatment in the medium term. This allows important parameters	See section 4.2.5 of the EAG report for the EAG's perspective regarding time horizon: "The main concerns of the EAG relate to the model time horizon of 12 months. The company indicated (response to clarification question B3a3) that [redacted] remain on treatment at the end of the 12-month time horizon. Therefore, it is questionable whether all relevant costs and benefits are captured within this period."

		(including dropout) to be observed as well as to demonstrate that treatment effect is maintained into the medium term. • A one-year timeframe is conservative from the perspective of estimated cost-effectiveness because it includes wasted prescriptions for those who do not take their medicine and that there is heterogeneity in treatment response. The desire for a longer-term model implied by the wording of this issue is simply that treatment may last for longer than one year. Despite the statement that this has not been appropriately modelled, the CS did include a longer-term model as a scenario rather than a base case. This is because, as demonstrated in the 301 and 303 studies, patients persisting with treatment to one year have better outcomes on average and based on the epidemiological relationship between poor sleep and poor long term health outcomes (increased risk of cardiovascular disease), the cost-effectiveness of long-term treatment improves on the one-year base case estimates presented.	
Other Issue 2: Extrapolation of short-term benefits and	No	As described in relation to Other Issue 1, a lifetime model of long-term daridorexant	Please see response to issue 7 above.

any potential waning of treatment effect over the long term (relating to Other Issue 1 and EAG Issue 7).		lowers the estimated ICER due to selective attrition and potential long term health benefits. There was no waning in the treatment effect on ISI© observed over 12 months and no reason (such as increased tolerance as with the benzodiazipenes and Z drugs) to expect this in the future. Therefore, there is no basis to estimate waning of treatment effect. In any case the importance of treatment waning in the future is most relevant to treatments which have stopped (for example – do treatment effects of a new cancer drug continue beyond the trial after treatment is discontinued?). For a once-a-day treatment like daridorexant, where the pharmacodynamics and pharmacokinetics support early onset of effect and early washout, any treatment waning should be noticed by the patient and should lead to treatment discontinuation – instigated either through the patient discontinuing themselves or by a treating physician at one of the recommended treatment review visits.	

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response	EAG comment

Additional issue 1: The EAG report notes uncertainty in the approach to exploring the impact of productivity losses as we explore two methods with different results, but does not sufficiently highlight the one directional positive impact and associated societal value of treating insomnia disorder.	Section 4.2.9.5	Yes	Whilst not raised as a specific issue we have included scenario analyses on indirect costs per "NICE health technology evaluations: the manual". Chronic insomnia impacts an individual's mental and physical health, quality of life (QoL), and productivity. Importantly, the consequences of insomnia go well beyond the individual, as there may be cascading effects on families, employers, and global economies. This is increasingly recognised, and the productivity impact is therefore a critical component of the overall cost-effectiveness of interventions for chronic insomnia. **There are certainly arguments that this should be the base case in our submission however we hope to make the case that the scenario with productivity included can be presented to the committee per section 4.4.23 of the NICE Manual on productivity costs - "They can be presented separately, as additional information for the committee, if such costs may be a critical component of the value of the technology."** When these costs are included directly using the SDS measured from the clinical trial, treatment with daridorexant is effectively cost neutral in year 1 and cost saving in subsequent years. When they are estimated indirectly from WPAI mapped to ISI©, daridorexant is cost saving after 3 months of treatment.	Thanks for this comment.

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response	EAG comment
			In our submission we reference RAND Europe "Why Sleep Matters: Quantifying the Economic Costs of Insufficient Sleep" There is now an additional report from RAND Europe [redacted] (RAND, 2023). We have permission to share the Executive Summary as a confidential document and this is attached as a reference. This report highlights [redacted]. This supports the broader case on the economic value of treating chronic insomnia.	

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)	EAG comment
Issue 14: There were several issues related to the mapping of ISI© scores to EQ-5D utilities, including the generalisability of the mapping function to the target sample, (lack of) conceptual overlap between ISI© and EQ-5D instruments, (lack of) validation of the mapping function	In the CS, the base case ICER estimated using the original ISI©-EQ5D mapping and gamma-log GLM was [redacted] / QALY.	As per the response to Issue 14, the company has incorporated the EAG's suggestions to improve the ISI©-EQ5D mapping. Additional modelling types were tested (CLAD and ALDVMM) and an additional validation exercise was undertaken. Although the suggested CLAD model performed poorly, the suggested ALDVMM model performed well – outperforming the original gamma-log GLM by a	Base-case ICER resulting from the change[redacted] / QALY Change from CS base-case ICER: [redacted] / QALY	See section 4.2.9 of the EAG report for the EAG comments on the estimated productivity losses

and (lack of) exploring other model types.		small margin. The CEM is now updated to include ALDVMM as a scenario which increases the base case ICER by a small amount.		

Sensitivity analyses around revised base case

PLEASE DESCRIBE HERE