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							

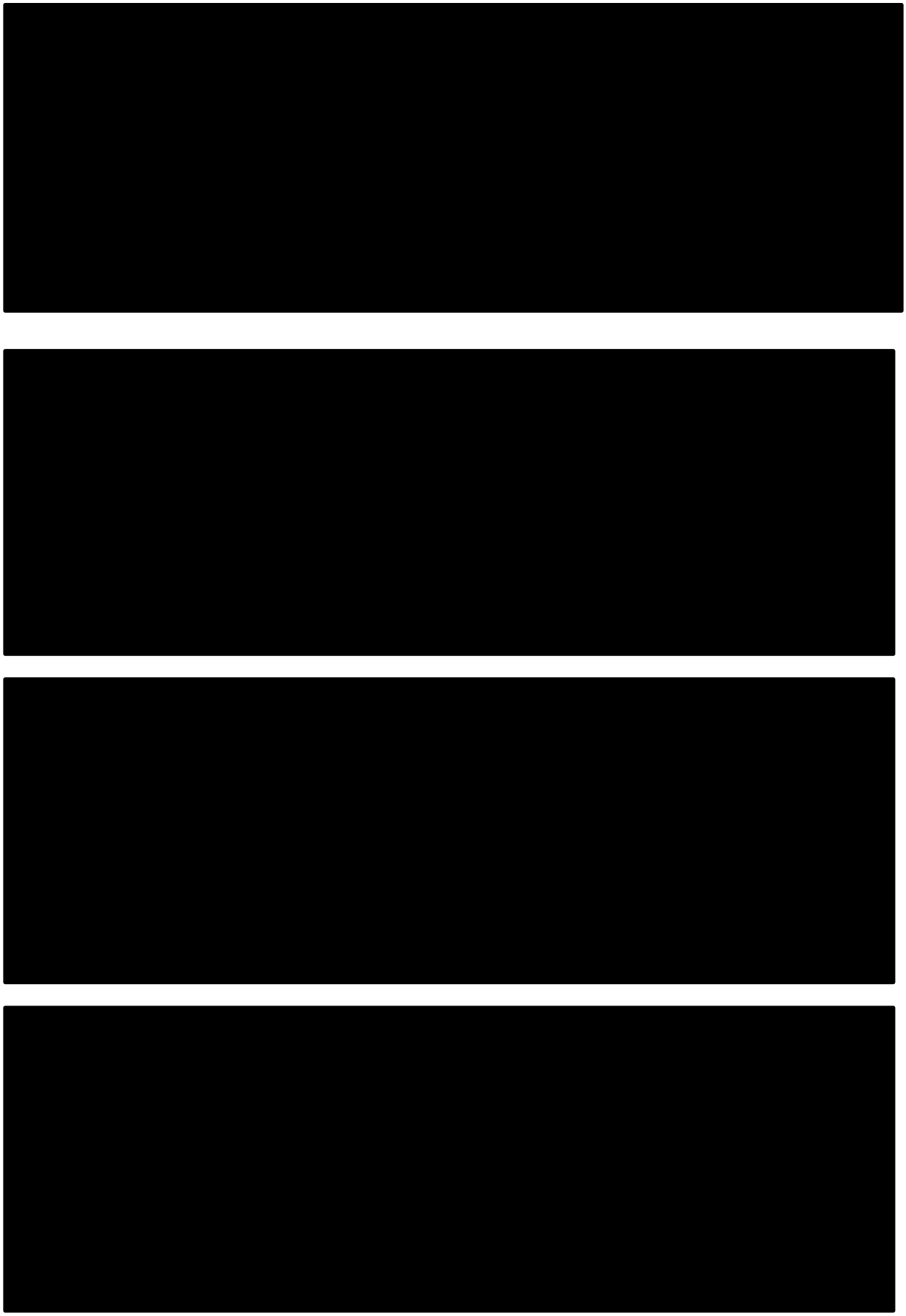
EAG comment:

- 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]

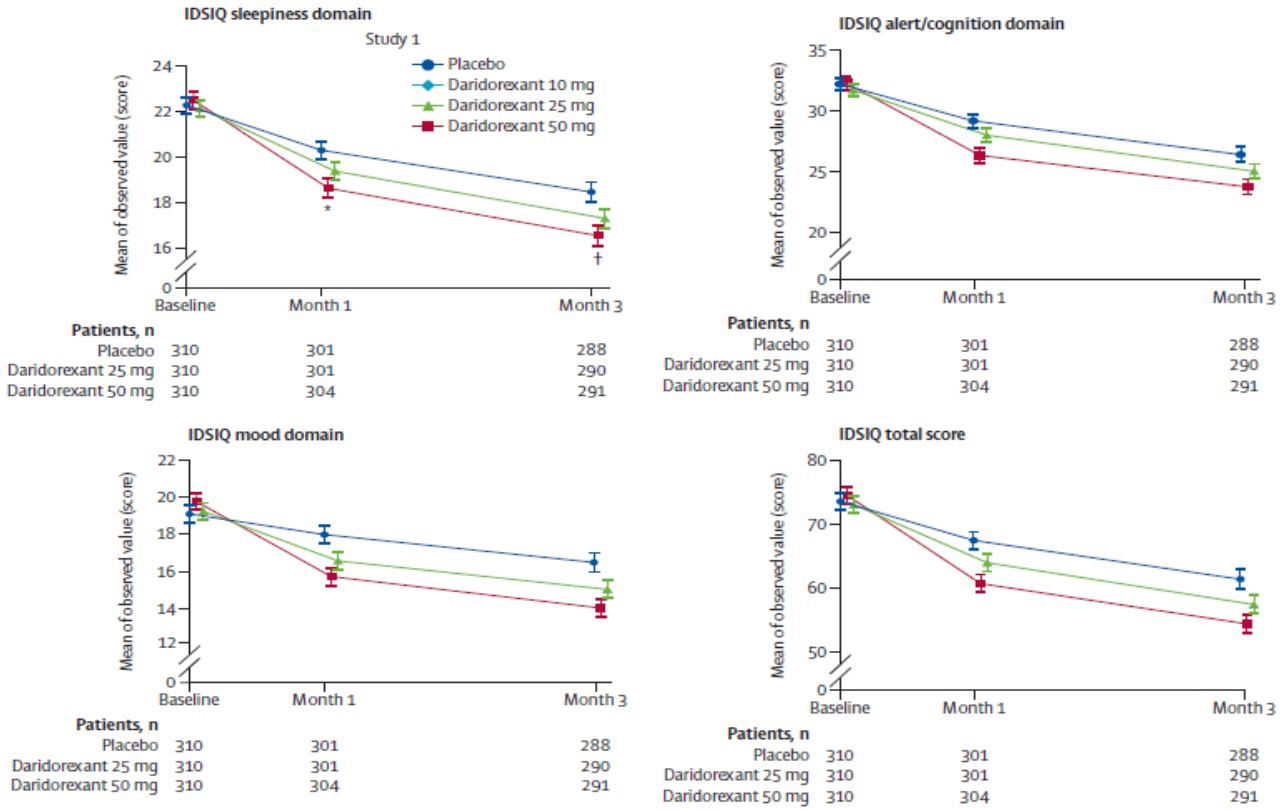
Subjects in the daridorexant 50 mg group reported significant reduction from baseline in IDSIQ sleepiness domain score compared to placebo at Month 3 (least squares mean (LSM) difference -1.90, [-2.95 to -0.98], p=0.0002) (Table 3.34).

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].

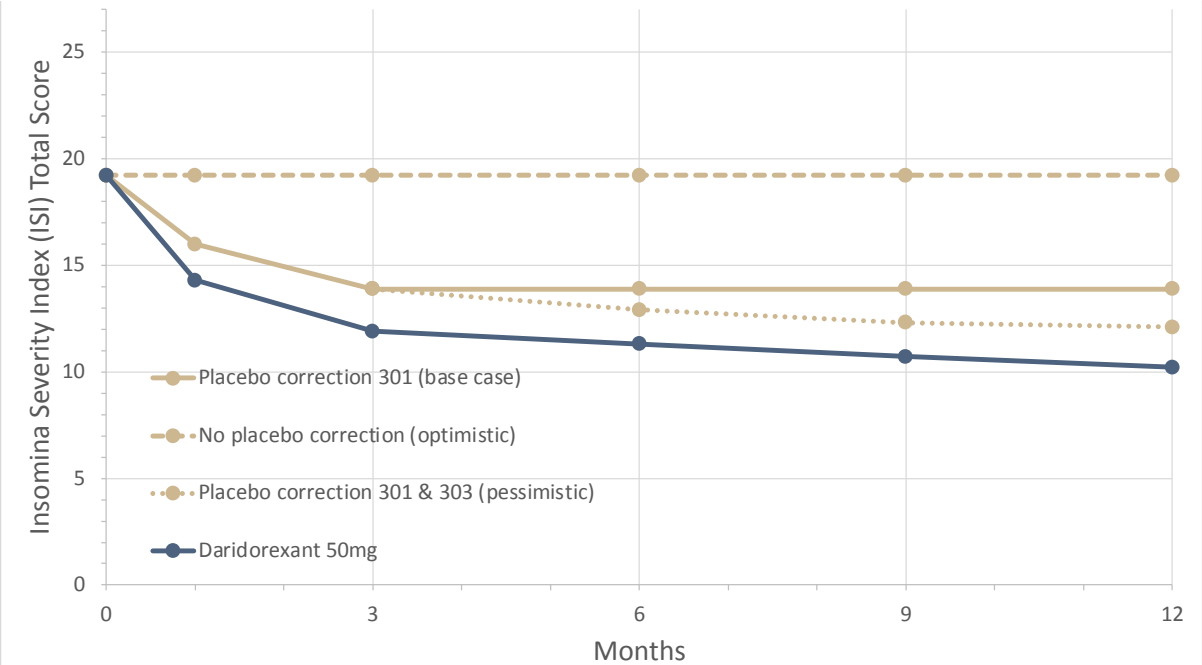
The effectiveness of daridorexant for the first 12 weeks was estimated via seemingly unrelated regression (SUR) of ISI© scores derived from Study 301 to adjust for both baseline ISI© and placebo effects. Although not clearly described in the CS, from the model file it appears that for the 6th, 9th and 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] 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.