NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Zuranolone for treating postnatal depression

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using zuranolone in the NHS in England. The evaluation committee has considered the evidence submitted by the company and the views of non-company stakeholders, clinical experts and patient experts.

This document has been prepared for consultation with the stakeholders. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the stakeholders for this evaluation and the public. This document should be read along with the evidence (see the committee papers).

The evaluation committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex or sexual orientation?

Issue date: August 2025

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The evaluation committee will meet again to consider the evidence, this evaluation consultation document and comments from the stakeholders.
- At that meeting, the committee will also consider comments made by people who are not stakeholders.
- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using zuranolone in the NHS in England.

For further details, see NICE's manual on health technology evaluations.

The key dates for this evaluation are:

- Closing date for comments: 18 September 2025
- Second evaluation committee meeting: 08 October 2025
- Details of membership of the evaluation committee are given in section 4.

1 Recommendations

1.1 Zuranolone should not be used to treat postnatal depression in adults.

1.2 This recommendation is not intended to affect treatment with zuranolone that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS healthcare professional consider it appropriate to stop.

What this means in practice

Zuranolone is not required to be funded and should not be used routinely in the NHS in England for the condition and population in the recommendations.

This is because there is not enough evidence to determine whether zuranolone is value for money in this population.

Why the committee made these recommendations

Usual treatment for postnatal depression is established clinical management. This includes antidepressant medicines, high-intensity psychological interventions (such as cognitive behavioural therapy) or both.

Clinical trial evidence suggests that, compared with placebo, zuranolone decreases the number of depressive symptoms. But this is uncertain because:

- the data was not collected over a long enough period, so the long-term benefit of zuranolone is uncertain
- there were differences in the populations of the trials from what would be expected in NHS practice
- established clinical management in the trials was different from what is offered in the NHS.

There is uncertainty in the economic model about:

- how long the benefit from treatment with zuranolone lasts
- how postnatal depression affects long-term quality of life
- to what extent other antidepressant use affects how well zuranolone works.

Because of the uncertainties in the clinical evidence and economic model, it is not possible to determine the most likely cost-effectiveness estimates for zuranolone. So, it should not be used.

2 Information about zuranolone

Anticipated marketing authorisation indication

2.1 Zuranolone (Zurzuvae, Biogen) does not have a marketing authorisation in Great Britain yet. The Committee for Medicinal Products for Human Use (CHMP) has adopted a positive opinion recommending the granting of a marketing authorisation for the medicinal product Zurzuvae, intended for 'the treatment of adults with postpartum depression'.

Dosage in the marketing authorisation

2.2 The dosage schedule will be available in the summary of product characteristics for zuranolone

Price

2.3 The list price for zuranolone is currently confidential until the UK marketing authorisation is granted.

2.4 The company has a commercial arrangement, which would have applied if zuranolone had been recommended.

Carbon Reduction Plan

2.5 Information on the Carbon Reduction Plan for UK carbon emissions for Biogen will be included here when guidance is published.

3 Committee discussion

The evaluation committee considered evidence submitted by Biogen, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the committee papers for full details of the evidence.

Postnatal depression

3.1 Postnatal depression is a form of depression that can occur after childbirth. This evaluation considers postnatal depression in people who have given birth. This can include women, trans men and non-binary adults (from here, referred to as 'people with postnatal depression'). The exact causes of postnatal depression are not fully understood, but it is thought that the pathophysiological factors driving it are genetic, metabolic, endocrine, environmental and neurological. Postnatal depression is one of the most common complications of pregnancy and childbirth. The Royal College of Psychiatrists has estimated that in 2024 between 10% and 15% of women who gave birth in England were diagnosed with postnatal depression.

Several factors may increase the likelihood of developing postnatal depression. These include a previous history of mental health issues and situational factors such as limited support from close friends or family, navigating a difficult relationship with a partner, recent significant life events such as bereavement, and physical or psychological stress. But postnatal depression can develop in the absence of any of these factors. Postnatal depression symptoms can start before birth (perinatal depression) and up to 12 months after giving birth. Common symptoms include consistently low mood, anxiety, heightened irritability towards a partner, baby or other children, persistent fatigue, difficulty sleeping, challenges with concentration or decision making, changes in appetite, recurring negative thoughts (including suicidal and infanticidal ideation), feelings of guilt, concerns about the baby's wellbeing, and difficulties with bonding or finding enjoyment in spending time with the baby. A patient

expert at the committee meeting explained that symptoms of postnatal depression can begin suddenly and severely. They explained that the symptoms, particularly anxiety, were debilitating and had an impact on day-to-day functioning. Postnatal depression can result in poor outcomes for the child, including cognitive delay, emotional disorders and behavioural problems. It can also put an additional strain on relationships with partners and families and can impact their quality of life.

Postnatal depression can be classified as mild, moderate or severe. In research, depression measurement tools, such as the 17-item Hamilton Rating Scale for Depression (HAM-D-17) and the Montgomery–Åsberg Depression Rating Scale (MADRS), are often used to assess the severity of postnatal depression and the efficacy of postnatal depression interventions. In clinical practice, the severity of an individual's postnatal depression will be based on clinical judgement of the intensity and number of symptoms they present with and the degree of functional disability. The clinical expert at the committee meeting explained that it is challenging to diagnose postnatal depression, particularly moderate or severe, in clinical practice. This is because an individual may have other severe mental health comorbidities, so it can be difficult to assess the severity of their symptoms of depression in isolation.

Postnatal depression often improves within a few months, but around 30% of people continue to have symptoms after a year. In 2020, suicide was the leading cause of deaths for people in the UK in the year after giving birth. The Avon longitudinal study of parents and children (ALSPAC; Boyd et al. 2013) found that women who had postnatal depression symptoms in the year after giving birth, especially those with persistent symptoms, were at increased risk of prolonged depression until at least 11 years after childbirth. The committee questioned whether having postnatal depression was a risk factor for a later diagnosis of major depressive disorder (MDD). Clinical advice to the EAG noted that after 1 year postnatal depression would be treated as MDD because the symptoms

are similar. The clinical expert at the committee meeting supported this. But the company argued that postnatal depression was distinct from MDD. The committee heard from the clinical expert that postnatal depression is likely to increase the risk of a subsequent diagnosis of MDD. The committee concluded that postnatal depression has a substantial impact on people after pregnancy, their partners and families.

Positioning of zuranolone

3.2 The population in this evaluation's scope was adults with postnatal depression. The company stated that zuranolone would be for people who had given birth (see section 3.1). The company submission noted that zuranolone will initially be prescribed by perinatal psychiatrists within community perinatal mental health services. So, it expects that it will most likely be used to treat moderate or severe postnatal depression. The EAG's clinical advisers agreed, noting that people treated in secondary NHS perinatal mental health services often have severe postnatal depression, complex postnatal depression or co-occurring psychiatric conditions. The College of Mental Health Pharmacy supported this, commenting that zuranolone would be more appropriate for treating severe postnatal depression and less appropriate for treating mild to moderate postnatal depression. The clinical expert at the committee meeting highlighted the difficulty of diagnosing someone with moderate or severe postnatal depression in a secondary perinatal mental health setting. This is because most people would present with other psychiatric comorbidities. They explained that a clinical diagnosis of moderate or severe postnatal depression would need to be made based on the symptoms of depression, not on the symptoms of their other comorbidities.

The positioning of zuranolone in secondary care was supported by the EAG's clinical advisers, NHS England and the clinical expert at the committee meeting. A submission from the College of Mental Health Pharmacy noted that zuranolone's rapid onset of action and potential

sedative effects would need specialist oversight for safe initiation and monitoring. It said that, unlike traditional antidepressant medicines managed in primary care, zuranolone would be used within perinatal psychiatry services to ensure appropriate assessment and diagnosis, risk management and follow up. NHS England and the clinical expert were cautious that a positive recommendation for zuranolone could lead to a sizeable increase in referrals to mother-and-baby units. The company indicated that people with postnatal depression would not need to be inpatients to have zuranolone. It highlighted that zuranolone's summary of product characteristics will say that it should be initiated by a specialist prescriber based on a referral to secondary care. The committee noted that it can only appraise zuranolone within its licence and cannot change this requirement. The patient experts acknowledged the benefits of a fast-acting treatment for postnatal depression but highlighted the unequal access to secondary perinatal mental health services across England. So, those who face barriers to accessing secondary care, such as people from ethnic minority backgrounds and those from low socioeconomic backgrounds, might have challenges accessing zuranolone (see section 3.16). The company recognised the need to support implementing a new treatment. But it expected that introducing this first postnatal depression-specific treatment would initiate a shift towards faster referrals into specialist care. The committee thought that the benefits of a short treatment course and the rapid onset of zuranolone could be limited by the delayed access to secondary care within the NHS, particularly for people who already face barriers to accessing secondary perinatal mental health services. The committee concluded that zuranolone would likely be used in secondary NHS perinatal mental health services to treat moderate to severe postnatal depression.

Clinical management

3.3 There are currently no licensed pharmacological treatments for postnatal depression in the UK. Treatment options are limited to established clinical management (ECM). NICE's guideline on the clinical management of

antenatal and postnatal mental health (CG192) states that the choice of treatment for postnatal depression depends on symptom severity and mental health history. For people with moderate or severe postnatal depression, CG192 recommends:

- high-intensity psychological interventions such as cognitive behavioural therapy
- pharmacological treatments.

Pharmacological treatments available in the NHS include tricyclic antidepressants (such as amitriptyline, clomipramine, dosulepin, imipramine, lofepramine and nortriptyline), selective serotonin reuptake inhibitors (such as fluoxetine, citalopram, escitalopram, paroxetine, sertraline) and selective norepinephrine reuptake inhibitors (such as duloxetine and venlafaxine). The clinical expert agreed that these medicines would be offered to people with moderate or severe postnatal depression. They noted that although antidepressant medicines are commonly prescribed to people with postnatal depression, they are not licensed for treating postnatal depression. They explained that current treatment options can take several weeks to improve symptoms of postnatal depression and need frequent follow ups to monitor effectiveness and side effects. This means that people with postnatal depression can be left with debilitating symptoms after birth. This was supported by the patient expert, who noted that it took 3 to 4 weeks for them to feel the benefit of an antidepressant, during which time they had difficulties with day-to-day functioning. The committee noted that ECM was the most appropriate comparator for zuranolone and concluded that a new fast-acting treatment specifically for postnatal depression would be welcome.

Clinical effectiveness

Data sources

3.4 The main clinical evidence to support zuranolone came from 2 randomised controlled trials: SKYLARK (n=200) and ROBIN (n=153). Both trials were phase 3, randomised, double-blind, placebo-controlled outpatient trials. Both trial populations included people aged 18 to 45 years with a major depressive episode that began during their third trimester or within 4 weeks of giving birth. Participants had severe postnatal depression, classified as a baseline HAM-D-17 total score of 26 or more. In both trials, participants were predominantly from the United States. People were excluded from the clinical trials for a broad range of reasons, including if they had comorbidities such as psychosis or if they had attempted or were at risk of suicide associated with the current episode of postnatal depression.

Both trials compared zuranolone with placebo, but the SKYLARK trial used the proposed license daily dose of 50 mg and the ROBIN trial used 30 mg. Zuranolone was taken for 14 days and participants were followed up to day 45 (4-week follow up). The SKYLARK trial permitted concomitant use of antidepressants and psychotherapy. The ROBIN trial permitted concomitant psychotropic medication. The clinical trials prespecified subgroups, including the use of antidepressants at baseline. The primary outcome measure in both trials was change from baseline in HAM-D-17 total score at day 15. Secondary outcomes included change from baseline in MADRS and Edinburgh Postnatal Depression Scale (EPDS) total scores from day 15, adverse events, change from baseline in Patient Health Questionnaire 9 (PHQ-9) and change from baseline in SF-36 score (ROBIN trial only). In its economic model, the company used the change from baseline EPDS score and safety data from the SKYLARK trial and the health-related quality-of-life (HRQoL) data collected through the SF-36 from the ROBIN trial.

Clinical trial results

3.5 A decrease in HAM-D-17 and EPDS score shows an improvement in depressive symptoms. The clinical trials found that zuranolone reduced the number of depressive symptoms compared with placebo. In the primary outcome, the SKYLARK trial found that the least-squares mean change from baseline in HAM-D-17 score at day 15 was statistically significantly greater for zuranolone than for placebo (-15.6 compared with -11.6, 95% confidence interval [CI] -6.3 to -1.7, p=0.001). The difference remained statistically significant until day 45 (-17.9 compared with -14.4, 95% CI -6.0 to -1.0, p=0.007). The least-squares mean change from baseline in EPDS total score in the SKYLARK trial showed a greater improvement in people having zuranolone compared with people having placebo at most timepoints, including day 15 (-2.0 difference, 95% CI -3.8 to -0.1, p=0.038) and day 45 (-2.4 difference, 95% CI -4.5 to -0.3, p=0.028). The clinical expert at the committee meeting indicated that the EPDS change from baseline results were at the lower end of what would be considered clinically meaningful. The committee noted some evidence of improvement in EPDS scores at other timepoints, but the 95% CI included the null value (exact values cannot be reported as the data is considered commercial in confidence), indicating that there may be no difference in effectiveness between zuranolone and the placebo treatment groups. The committee concluded that at day 15 zuranolone improved depressive symptoms compared with placebo. But it noted that the long-term effectiveness of zuranolone for treating postnatal depression was uncertain because of the short follow-up periods in the clinical trial evidence.

Generalisability

Clinical trial population

3.6 The clinical trials for zuranolone recruited people with severe postnatal depression, classified as a baseline HAM-D-17 score of 26 or more (see section 3.4). The EAG noted that the clinical trial populations did not align

with the scope population (adults with postnatal depression). It stated that the clinical evidence for zuranolone is mostly applicable to severe postnatal depression. This was supported by a clinical expert at the committee meeting. The company said that the clinical evidence for zuranolone provides data on the efficacy for moderate and severe postnatal depression. This is because when postnatal depression severity was measured using MADRS instead of HAM-D-17, people in the trial were classed as having either moderate or severe postnatal depression. The EAG highlighted that the mean baseline MADRS score for those with moderate postnatal depression suggests that it includes people at the upper end of the moderate severity classification. So, the evidence for the moderate postnatal depression population proposed by the company is not representative of a moderate postnatal depression subgroup in NHS clinical practice. The committee questioned what diagnosis this trial population would receive in clinical practice and whether it would align with the NHS population expected to have zuranolone (see section 3.2). The clinical expert reiterated that these depression measurement tools are not used to diagnose postnatal depression in clinical practice (see section 3.1). The company highlighted that a post-hoc analysis by postnatal depression severity using the MADRS classification tool found that there was no difference in treatment effect size between people with moderate postnatal depression and people with severe postnatal depression. The EAG thought that, based on the clinical trial data, the efficacy of zuranolone for mild or moderate postnatal depression could not be determined. The committee acknowledged that the clinical trial population was mostly people with severe postnatal depression. It noted that defining moderate and severe postnatal depression in clinical practice is challenging and recalled that people treated in secondary care, where zuranolone would be available, were likely to have a more severe postnatal depression diagnosis (see section 3.2). So, the committee thought that the severity of the clinical trial population was likely to align with the expected population in NHS clinical practice.

The clinical trials also excluded people who had comorbidities, had attempted suicide or were at risk of suicide during the current episode of postnatal depression, or had a history of substance abuse (see section 3.4). The EAG was concerned that these exclusions limited the generalisability of the trial results to the NHS target population. But it acknowledged that such exclusions are common in clinical trials. This was supported by submissions from clinical experts, who commented that a large proportion of people with postnatal depression who are treated in secondary care have comorbidities (see section 3.2). The clinical expert at the committee meeting said that they expected zuranolone to have a smaller effect on those people excluded from the clinical trial, suggesting that having comorbidities could be a treatment effect modifier. The EAG also noted that the majority of the trial population was from the United States, which has a different care system and ECM treatments available. The impact this could have on the clinical trial results was unknown. But the clinical expert told the committee that it did not expect the differences between a US and UK population to have a big impact. The committee noted that some differences between the clinical trial population and NHS population could have had a modifying effect on the clinical trial results. It concluded that there was uncertainty in the generalisability of the clinical trial population to the licensed population.

Established clinical management definition

3.7 ECM in the clinical trials was the use of antidepressant medicines and non-pharmacological interventions in both treatment arms (see section 3.4). The EAG noted that this definition excluded several treatments identified as ECM in the scope of this appraisal. Importantly, this excluded treatments that might be used by those with severe postnatal depression, such as augmentation (combinations of treatments, such as antidepressants from different drug classes). During the SKYLARK trial, people were not permitted to change treatments and had to wait until after the trial to access other ECM treatments. The EAG was concerned that these trial restrictions may have affected the outcomes, particularly for

people in the placebo group because they could not access the full range of treatment options that would be available in clinical practice. The company stated that the concomitant treatments permitted in the SKYLARK trial were expected to reflect NHS clinical practice. But a clinical expert submission commented that ECM in the trials was not comparable to what is offered in the NHS and they would not expect the results to be generalisable to the UK.

The EAG was also concerned about the comparability of the observed use of antidepressant medicines and psychotherapies (not reported) in the trials compared with the NHS population. In the SKYLARK trial, 15% of people in each treatment group used an antidepressant medicine at baseline (see section 3.4). The EAG's clinical advisers noted that most of the people (more than 90%) eligible for zuranolone in clinical practice would already be taking an antidepressant. So, the trial population may not fully represent those considered for treatment with zuranolone in the NHS. The company explained that around 85% of the SKYLARK trial population had zuranolone as a monotherapy. It stated that zuranolone is expected to be used with or without ECM, regardless of history of prior antidepressant use. The clinical expert at the committee meeting noted that in most clinical scenarios zuranolone would be prescribed in addition to ECM. They estimated that about 70% of people in perinatal mental health care would be having an antidepressant and they expected that the treatment effect of zuranolone would be smaller for people having an antidepressant. This aligned with the clinical trial results for SKYLARK that found that the least-squares mean change from baseline in HAM-D-17 favoured placebo for people with baseline antidepressant use. But this analysis was based on a small proportion of people and was not powered to detect a difference in efficacy. The committee agreed that there were differences in the definition of ECM in the clinical trial compared with NHS clinical practice and that ECM use could potentially be a treatment effect modifier. It concluded that there was uncertainty in the generalisability of the clinical trial results to the target NHS population.

Economic model

Company's modelling approach

3.8 The company used a patient-level simulation model to track postnatal depression (based on EPDS score), costs and health outcomes of people having zuranolone with ECM or ECM alone. The model used a 29-year time horizon and assumed that people started treatment 5 months after giving birth, based on the average time of treatment initiation in the SKYLARK trial. The model simulated individual patient characteristics (age, baseline EPDS score and antidepressant use) and calculated the treatment effects in the model using EPDS scores from SKYLARK in the short term and the ALSPAC data in the long term (see section 3.10). The model had 5 EPDS health states (0, 1 to 6, 7 to 13, 14 to 18 and 19 to 30) and people were classified as being either in remission (EPDS score of less than 13) or not in remission (EPDS score of 13 or more). The costs applied varied based on whether an individual was in remission. Health outcomes were measured in quality-adjusted life years (QALYs), which were calculated based on utilities associated with each of the EPDS health states using SF-36 data collected during the ROBIN trial (see sections 3.11 and 3.12). The EAG agreed that the model structure used was appropriate because it was able to capture changes in EPDS score over time as a continuous variable and could account for non-linear effects of baseline characteristics on patient EPDS trajectories. But, the EAG was concerned that the model estimated long-term cost and health effects over the 29-year time horizon based on immature trial data collected over only 45 days. It thought that the clinical trials were too short to show how a change in EPDS score relates to changes in postnatal depression severity or the risk of developing MDD in the longer term. Also, the trials were too short to inform what proportion of people per EPDS category would go on to develop MDD after postnatal depression. The EAG also noted that the cost of increased referrals to secondary care predicted by the company if zuranolone is available were not included in the model (see section 3.2). The committee thought that these costs

should be included in the model. The committee recalled that postnatal depression was likely to increase the risk for a subsequent diagnosis of MDD (see section 3.1).The committee concluded that the model was suitable for decision making, but acknowledged the uncertainty about the duration of the treatment effect of zuranolone (see section 3.12) and the long-term utility estimates (see section 3.13) because the data to inform the model came from trials with a very short follow-up duration.

Modelled population

3.9 The company stated that the modelled population was adults with postnatal depression. The committee was aware that the population most likely to have zuranolone would have moderate to severe postnatal depression (see section 3.2) and that the evidence informing the model was largely for severe postnatal depression (see section 3.6). The company explained that it could not provide separate subgroup analyses for the moderate and severe subgroups because it did not have long-term EPDS trajectory data by severity. To explore the impact of severity, the EAG ran 2 scenario analyses that included results from only simulated people with a baseline EPDS score of either more than 19 (severe postnatal depression) or less than 19. It chose to use people with severe postnatal depression in its base case. The committee noted that the population modelled did not have a large impact on the incremental cost-effectiveness ration (ICER). It recalled its conclusion that the severity of postnatal depression in the people included in the clinical trials would probably align with those who would receive zuranolone in clinical practice (see section 3.6). So, the committee concluded that the full SKYLARK trial population should be modelled.

Treatment effect duration

3.10 In the company's model, treatment effectiveness is informed by changes in EPDS from SKYLARK (see section 3.5). The 45-day follow-up data was extrapolated to estimate the short-term EPDS score at 3 months post-treatment initiation in each treatment arm. The difference in treatment

effect between the treatment arms at 3 months was assumed to be maintained for 29 years (the full time horizon) by the company. The treatment effect in the long term was estimated using the ALSPAC long-term data to inform the EPDS trajectory of the zuranolone with ECM arm. Long-term EPDS in the ECM arm was then estimated by applying a fixed increase in EPDS score, which maintained the treatment effect benefit of zuranolone with ECM compared with ECM alone found at the end of the short-term model (3 months after treatment initiation). The committee noted that the model did not account for the uncertainty in the EPDS score measured by the 95% CI (see section 3.5). It stated that this uncertainty could be better captured in the model using a Bayesian framework with a distribution of EPDS scores.

The EAG noted that people in the ALSPAC study had only ECM and not zuranolone, so the data could not inform the long-term treatment effect of zuranolone. So, the EAG thought it was more appropriate to use the results from the ALSPAC study to inform long-term EPDS scores in the ECM arm but not the zuranolone with ECM arm. For this arm, the EAG estimated the EPDS scores by adding the zuranolone treatment effect to the ECM arm scores. The EAG was also concerned about substantial data missing in ALSPAC that may not have been adjusted for. The committee noted that both approaches to informing the long-term EPDS scores had a small impact on the ICER. It concluded that it was most appropriate to use the ALSPAC data to inform the long-term EPDS trajectory in the ECM arm. But it acknowledged that the level of missing data in ALSPAC added to the uncertainty in the long-term estimates of EPDS scores.

The EAG was concerned about immature trial evidence being used to inform long-term model assumptions. It noted that assuming the treatment benefit from a 14-day treatment course would be maintained for 29 years was implausible. This is because there is no data available on the treatment efficacy of zuranolone beyond 45 days in postnatal depression.

The EAG highlighted that extrapolations of the treatment effect are highly uncertain because of the limited long-term data available. The EAG preferred to include a waning of zuranolone's treatment effect over time. With waning, the long-term EPDS curves in both treatment arms will eventually converge. In response to clarification, the company provided a scenario that assumed that people in both treatment arms would eventually be in remission. This was based on the assumption that the NHS would intervene to support people with postnatal depression to reach remission in the long term. To model this, people with an EPDS of 13 or more (not in remission) at a fixed timepoint (2, 5 or 10 years) in both treatment arms had their EPDS scores gradually decreased to 12.99 (in remission) over a 24-month period. The EAG provided 7 alternative treatment effect waning scenarios:

- zuranolone's treatment benefit starts to wane 3 months after treatment initiation and the EPDS curves from both treatment arms converge after a further 6 months, 1 year, 2 years, 5 years or 10 years
- zuranolone's treatment benefit starts to wane 1 year after treatment initiation and the EPDS curves from both treatment arms converge after a further 6 months or 1 year.

The EAG chose for its base case to wane the treatment benefit 1 year after treatment initiation and for the EPDS curves to converge after a further year. This is because it acknowledged that the treatment benefit of zuranolone may continue beyond the postnatal depression period (1 year after giving birth) but noted that it would not continue for an implausibly long time. The clinical expert at the committee meeting agreed with the EAG that the long-term EPDS trajectory curves would eventually converge. They estimated that, based on research outside of postnatal depression and unrestricted by severity, the treatment benefit would disappear after 1 year. The committee considered the treatment effect waning scenarios presented and recalled that zuranolone was given for a 14-day treatment course. It decided that there was

insufficient clinical evidence available to support a long-term treatment effect benefit for zuranolone. So, it agreed that treatment waning should be applied. The committee noted that there was no evidence presented to support the more optimistic waning scenarios (where the benefit of zuranolone lasts 2 years or more). But it thought that there was clinical plausibility in the other scenarios presented. The committee concluded that, based on the clinical expert feedback, it was also plausible that the treatment effect could start to wane 6 months after treatment initiation and last for 6 months.

Utility values

Choice of tool to estimate utilities

3.11 For each of the 5 EPDS health states in the company's model, HRQoL was informed by SF-6D utility values derived from SF-36 data collected in the ROBIN trial (see section 3.4). These were estimated by pooling the SF-36 data from both the zuranolone and placebo arms in the ROBIN trial for the same EPDS category. The estimated utility values were applied for the entire time horizon (29 years). The EAG was concerned about applying utilities based on 45 days of data for the duration of the time horizon. It also highlighted the relatively small sample size of the ROBIN trial, so some utility values were likely to be estimated from very few people.

The NICE reference case recommends using utilities estimated from the EQ-5D tool (see section 4.3.12 of NICE's manual on health technology evaluations). It states that if EQ-5D values are not available, utilities should be estimated from another measure using statistical mapping or evidence should be provided to show that the EQ-5D is not appropriate to use. The zuranolone clinical trials did not collect EQ-5D-reported utilities. The company provided a scenario analysis that used utilities mapped from the PHQ-9 values collected in the ROBIN trial to the EQ-5D. To apply these in the model, the scenario analysis used 10 EPDS categories: 0, 1 to 3, 4 to 6, 7 to 9, 10 to 12, 13 to 15, 16 to 18, 19 to 21, 22 to 24 and 25

to 30. The company preferred to use the SF-36 mapped to SF-6D values. This is because the company felt that it captures the broader aspects of quality of life that are affected by postnatal depression, such as physical and social functioning. The EAG disagreed and preferred to use the scenario of PHQ-9 mapped to EQ-5D utilities in its base case. This is because the EQ-5D has a specific depression and anxiety domain and the PHQ-9 is designed and validated to measure depression symptoms and responsiveness. It also noted that the SF-36 is not NICE's preferred utility tool and produces different utilities from the EQ-5D, and using it would reduce the comparability of QALYs across conditions. The company acknowledged that using the EQ-5D-mapped utilities would be an acceptable approach. The committee noted that the EQ-5D was NICE's preferred measure of HRQoL. It also noted that there were implausible differences between EPDS health-state utilities when measured by the SF-6D, which were resolved by using EQ-5D utilities. The committee concluded that the EQ-5D utilities were the most appropriate utilities to model postnatal depression.

Long-term utility values

3.12 The EAG thought it was appropriate to use different utilities to model the postnatal depression period (EQ-5D utilities applied up to 1 year after treatment initiation) and the post-postnatal depression period. This is because of clinical advice that postnatal depression lasts for 1 year after birth and would then be treated as MDD (see section 3.1). It noted that the company's systematic literature review did not identify any long-term postnatal depression utility values, so the EAG used MDD utilities to inform changes in quality of life after the postnatal depression period. EQ-5D utilities from NICE's technology appraisal guidance on vortioxetine for treating major depressive episodes (TA367) were used to inform utility values in the 'remission' (0.85) and 'non-remission' (0.67) health states. It explained that this aligned with how costs were modelled (see section 3.8). The EAG acknowledged that there is not a clear correlation between a clinician's definition of 'remission' and the definition used in the model

(an EPDS score of 13 or more). In its base case, the EAG applied the TA367 utilities 1 year after treatment initiation to align with treatment effect waning in the model (see section 3.10). The company disagreed with using MDD utilities after the first year. It explained that modelling utilities by remission status rather than EPDS score assumes that everyone with an EPDS score of 13 and over (not in remission) has the same quality of life. The company highlighted that someone with an EPDS score of 30 would probably have a very different quality of life from someone with an EPDS score of 15 who is almost in remission. The committee thought that after the first year postnatal depression would probably be treated as MDD. It also acknowledged that it would be preferable to have more granularity of MDD utilities applied for people not in remission, to reflect the difference in EPDS score seen in the postnatal depression data. The committee requested analyses that explored using MDD utility data in the post-postnatal depression period. This includes modelling MDD utilities with greater granularity to account for changes in quality of life across EPDS scores, for example finding MDD utilities by HAM-D-17 and then mapping these to EPDS scores in the model.

Concomitant antidepressant use

3.13 ECM in the company's model included the cost of concomitant antidepressant medicine (sertraline or fluoxetine only). The company assumed 15.3% concomitant antidepressant use in both treatment arms of the modelled population, as per the SKYLARK trial. The EAG thought that the proportion of people using an antidepressant is likely to be much higher than 15.3%, based on clinical expert advice that most people in perinatal mental health care would be having an antidepressant (see section 3.7). The EAG appreciated that the impact of baseline antidepressant use on efficacy could not be explored in the model because the SKYLARK trial sample size was too small to estimate a subgroup effect. Instead, the company provided a scenario which assumed that 100% of people incurred the cost of baseline antidepressant use in the short-term EPDS prediction model. The EAG ran additional

scenarios that assumed 50%, 80% or 90% of people incurred the costs of having baseline antidepressant medicines. The EAG used the 90% baseline antidepressant use scenario as its base-case assumption. Both the company's and the EAG's scenario analyses slightly increased the ICER. The EAG noted that these scenario analyses did not capture the potential interaction between antidepressant use and the zuranolone treatment effect (see section 3.7), so the true impact on cost effectiveness is unknown. The committee recalled its conclusion that 70% of people would probably be having an antidepressant medicine in perinatal mental health care and that antidepressant use was likely to be a treatment effect modifier (see section 3.7). So, the committee concluded that 70% of people in the model should be assumed to have baseline antidepressant use. It said that it would like the company to explore the impact of antidepressant use being a treatment effect modifier. It noted that this could include using meta-analysed data from the ROBIN and SKYLARK trials to estimate a subgroup effect or looking at data from other data sources, such as the MOUNTAIN clinical trial.

Mortality

3.14 The company's model assumed that zuranolone had no impact on the risk of mortality for people with postnatal depression because there was no evidence collected to support this assumption. So, it applied general population mortality rates equally to both treatment arms in its base case. The company ran a scenario analysis that applied a standardised mortality ratio (SMR) of 2 to people with an EPDS score of 13 or more (not in remission) in both treatment arms. This was based on a systematic literature review that looked at the risk of mortality for people with MDD compared with the general population. The College of Mental Health Pharmacy submission commented that zuranolone is unlikely to directly increase length of life more than established care, as it is specifically designed to address severe postnatal depression rather than have a direct impact on mortality. The EAG agreed that there is no evidence available of a direct effect of zuranolone on mortality, but it thought it was

reasonable that there may be an indirect effect through changes in EPDS scores. So, the EAG applied an SMR of 2 for people not in remission in its base-case assumptions. The committee acknowledged that suicide is the leading cause of death for people in the year after giving birth (see section 3.1). It noted that data on the impact of zuranolone on mortality had not been collected and that the clinical trial excluded people who may be at the highest risk of mortality (see section 3.6). But it was minded to accept that an impact on mortality should be modelled. The committee concluded that it was appropriate to model a mortality benefit. In the absence of an evidence-based estimate, the assumption of an SMR of 2 for people not in remission postnatal depression was reasonable.

Other factors

Uncaptured benefits and impacts

3.15 The committee considered whether there were any uncaptured benefits or disadvantages of postnatal depression or zuranolone treatment within the model. Stakeholders raised several impacts of postnatal depression and zuranolone that were not accounted for in the model. This included the impact of postnatal depression on children, partners and family members. NHS England highlighted many unknowns about zuranolone. These included the safety of repeated treatment courses, the risks of dependence or abuse and the potential impact of side effects such as sedation on the person taking zuranolone and their baby. This concern was shared by the patient experts at the committee meeting. They noted that if there was a sedating effect, people taking zuranolone would be likely to need additional support from partners, friends or family to complete daily tasks and care for their baby. They highlighted that there could be negative consequences for people who do not have access to this support, such as single parents. For this reason, people may prioritise taking a treatment without sedative effects. The clinical and patient experts were also concerned that the safety of zuranolone during breastfeeding had not been established. The patient expert explained the importance to some people with postnatal depression to continue

breastfeeding, so this may be a key consideration in treatment choice. Stakeholders also noted that the model does not fully capture all the potential benefits of zuranolone, including the subsequent positive impact that the fast-acting symptom relief could have for the person's baby, partner and family (see section 3.2). The committee acknowledged these uncaptured benefits and consequences. It concluded that it would take into account any potential uncaptured benefits and disadvantages when it was presented with responses to its requests for additional analyses.

Equality

3.16 The committee discussed the potential equality issues raised by stakeholders. These included people from ethnic minority backgrounds in the UK being at a greater risk of developing postnatal depression and being less likely to initiate treatment, access community mental health services and have follow-up care for postnatal mental health illness. Race is a protected characteristic under the Equality Act 2010. The committee recognised that pregnancy and maternity is also a protected characteristic, which includes everyone in the evaluation population. Stakeholders noted that there are challenges in current NHS clinical practice with diagnosing postnatal depression and unequal access to specialist perinatal mental health services across England. For example, people from lower socioeconomic backgrounds, people from ethnic minority backgrounds and people who live in remote areas may have limited access to the full services offered by the NHS. The committee was aware that if zuranolone was recommended, it would be available only in secondary perinatal mental health specialist services (see section 3.2). It noted that it can only appraise zuranolone within its licence and cannot change this requirement. Although people from ethnic minority backgrounds may be at greater risk of developing postnatal depression, this cannot be resolved by technology appraisal guidance. The committee stated that socioeconomic background and geographical access were not equalities issues within the scope of the Equality Act 2010. Geographical access to services is an implementation issue that cannot be addressed

by a NICE technology appraisal. The committee noted that its recommendation does not restrict access to treatment for some people over others. The committee concluded that the impact of zuranolone being prescribed only in secondary perinatal mental health specialist services is a potential health inequality (see section 3.17).

Health inequality

3.17 The committee acknowledged that a recommendation for zuranolone could improve overall population health, but it may also exacerbate the health inequality for people who have challenges accessing secondary perinatal mental health specialist services (see section 3.16). The committee noted that a NICE recommendation does not include guidance on service delivery or supporting implementation for disadvantaged groups. The committee noted that it can only recommend technologies as an option for use in NHS. Differences in uptake may determine health inequality impacts and be relevant to the committee's deliberations, but they cannot be addressed by the committee's recommendations (see section 4.12.5 of NICE's manual on health technology evaluations). The committee thought that it could understand this issue better and appropriately consider it within its deliberations, if health inequality evidence had been provided for the potential impact of a positive recommendation for zuranolone (as in sections 3.3.29 and 4.12 of NICE's manual on health technology evaluations). The committee noted that prescribing zuranolone in only secondary perinatal mental health specialist services has the potential to exacerbate these health inequalities. If recommended, consideration of potential difficulties in uptake in people from lower socioeconomic backgrounds, people from ethnic minority backgrounds and people who live in remote areas will be particularly important to ensure that existing health inequalities are not exacerbated. The committee noted that this would also help to ensure that population health gains are maximised. But it noted that this is beyond its remit.

Cost-effectiveness estimates

Company and EAG cost-effectiveness estimates

3.18 The company's and the EAG's base cases differed because of the following assumptions:

- modelled population (see section 3.9)
- parameterisation of long-term EPDS model (see section 3.10)
- treatment effect duration (see section 3.10)
- source and modelling of utility values for the postnatal depression period and post-postnatal depression period (see sections 3.11 and 3.12)
- the proportion of people having antidepressant medicine at the model's baseline (see section 3.13)
- modelling of postnatal depression mortality (see section 3.14)

The company's probabilistic base case was £12,706 per QALY gained. The EAG's probabilistic base case was £188,102. The key drivers of the ICER in the EAG's base case are the introduction of treatment effect waning (see section 3.10) and the application of MDD utilities in the long term after the postnatal depression period (see section 3.12).

Acceptable ICER

3.19 NICE's manual on health technology evaluations notes that, above a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of a technology as an effective use of NHS resources will take into account the degree of certainty around the ICER. The committee will be more cautious about recommending a technology if it is less certain about the ICERs presented. But it will also take into account other aspects, including uncaptured health benefits. The committee noted the high level of uncertainty, specifically that:

- there is no data available on the long-term treatment effectiveness of zuranolone in postnatal depression (see sections 3.4, 3.5 and 3.10).

- the clinical trial population had more severe postnatal depression than the licensed population (see section 3.6)
- aspects of the clinical trials lacked generalisability to NHS clinical practice and these differences could be treatment effect modifiers (see sections 3.6 and 3.7)
- there was no data collected about the impact of zuranolone on the risk of mortality (see section 3.14)
- there is no data available to inform long-term HRQoL in people with postnatal depression after the first year (see section 3.12)
- antidepressant use is a potential treatment effect modifier that is not accounted for in the model (see sections 3.7 and 3.13).

So, the committee concluded that further analyses are needed from the company before it can determine an acceptable ICER.

Committee's preferred assumptions

3.20 The committee's preferred assumptions were to model:

- the whole SKYLARK population (see section 3.9)
- the long-term EPDS trajectory of the ECM arm using the ALSPAC data and apply the zuranolone treatment effect to estimate the zuranolone with ECM arm (see section 3.10)
- treatment effect waning – based on clinical expert opinion, this could plausibly start 6 months after treatment initiation and last for 6 months (see section 3.10)
- EQ-5D utilities applied during the postnatal depression period and MDD utilities applied during the post-postnatal depression period (see sections 3.11 and 3.12)
- 70% of people having antidepressant medicine at baseline (see section 3.13)
- an increased risk of mortality for people with postnatal depression using an SMR of 2 for those who were not in remission (classified as an EPDS of 13 or more; see section 3.14).

The committee noted that a preferred ICER could not be determined and it would need additional information to help resolve some of the remaining uncertainties. The committee would like to see:

- An exploration of whether having comorbidities and concomitant ECM, in particular antidepressant medicines, is a treatment effect modifier for zuranolone (see sections 3.6 and 3.7). Where possible, the interaction with antidepressant use should be accounted for within the model (see section 3.13).
- The cost of increased secondary care referrals added into the model (see section 3.8).
- A scenario applying a Bayesian framework that uses a distribution of EPDS scores to account for uncertainty in the clinical trial results (see sections 3.5 and 3.10).
- Post-postnatal depression utilities informed by MDD utility data but applied with greater granularity to account for changes in quality of life across EPDS scores (see section 3.11).

Conclusion

Recommendation

3.21 The committee recalled the uncertainties it identified with the company's clinical-effectiveness evidence and economic modelling assumptions. It agreed that more exploration was needed around concomitant antidepressant use as a potential treatment effect modifier and the modelling of treatment effect waning and utilities before it could decide the most appropriate cost-effectiveness estimate. So, it could not recommend zuranolone for treating postnatal depression.

4 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee B.

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The minutes of each evaluation committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Charles Crawley
Chair, technology appraisal committee B committee

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser, a project manager and an associate director.

Cara Gibbons
Technical lead

Rachel Williams
Technical adviser

Jeremy Powell
Project manager

Emily Crowe
Associate director

ISBN: [to be added at publication]