NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Vortioxetine for treating major depressive episodes

The Department of Health has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using vortioxetine in the NHS in England. The Appraisal Committee has considered the evidence submitted by the company and the views of non-company consultees and commentators, and clinical experts and patient experts.

This document has been prepared for consultation with the consultees. It summarises the evidence and views that have been considered, and sets out the draft recommendations made by the Committee. NICE invites comments from the consultees and commentators for this appraisal (see section 9) and the public. This document should be read along with the evidence base (the Committee papers).

The Appraisal 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 provisional 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 race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

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 Appraisal Committee will meet again to consider the evidence, this appraisal consultation document and comments from the consultees.
- At that meeting, the Committee will also consider comments made by people who are not consultees.
- After considering these comments, the Committee will prepare the final appraisal determination (FAD).
- Subject to any appeal by consultees, the FAD may be used as the basis for NICE's guidance on using vortioxetine in the NHS in England.

For further details, see the Guides to the technology appraisal process.

The key dates for this appraisal are:

Closing date for comments: 1 July 2015

Second Appraisal Committee meeting: 9 July 2015

Details of membership of the Appraisal Committee are given in section 8, and a list of the sources of evidence used in the preparation of this document is given in section 9.

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

1 Appraisal Committee's preliminary recommendations

1.1 The Committee is minded not to recommend vortioxetine within its marketing authorisation, that is, for treating major depressive episodes in adults.

1.2 The Committee recommends that NICE requests further clarification and analyses from the company, which should be made available for the second Appraisal Committee meeting, and should include:

- a cost-effectiveness analysis of vortioxetine compared with relevant treatment options third line and beyond in the treatment pathway for major depressive disorder, for example, after selective serotonin reuptake inhibitors and 1 serotonin–norepinephrine reuptake inhibitor (venlafaxine), or as an option in a secondary care setting, or in patients who have had multiple adverse reactions
- consideration of these subgroups and to include probabilistic analyses that:
  - incorporate the broader evidence base for antidepressants, including at first-line treatment
  - define treatment success, and decisions to switch treatment, by remission and response
  - use the time point in which patients change to another treatment from the trials for the time point in the model (for example, 8 weeks rather than 4 weeks)

  - consider that people may receive treatment for up to 2 years (for example, to consolidate response)
  - include a risk of relapse at all stages of depression
  - use utility values from REVIVE
  - include a 24-month time horizon
  - present pairwise comparisons and incremental analyses for the probabilistic cost-effectiveness estimates
- disaggregated results for each of the pairwise comparisons.

2 The technology

2.1 Vortioxetine (Brintellix, Lundbeck) is an antidepressant with several modes of action that is thought to exhibit its clinical effect through direct modulation of receptor activity and inhibition of the serotonin transporter. Vortioxetine has a marketing authorisation in the UK "for the treatment of major depressive episodes in adults".

2.2 The summary of product characteristics lists the following 'common' and 'very common' adverse reactions for vortioxetine: abnormal dreams, constipation, decreased appetite, diarrhoea, dizziness, generalised itching, nausea and vomiting. For full details of adverse reactions and contraindications, see the summary of product characteristics.

2.3 Vortioxetine is administered orally. The recommended dosage is 10 mg once daily in adults younger than 65 years, and 5 mg once daily in adults 65 years and older. Depending on how the symptoms responds in an individual patient, the dose may be increased to a maximum of 20 mg once daily or decreased to a minimum of 5 mg once daily. Treatment for at least 6 months is recommended after the symptoms resolve. The price of a pack (28 tablets) of 5 mg, 10 mg or 20 mg tablets is £27.72 (excluding

VAT; company's submission). Costs may vary in different settings because of negotiated procurement discounts.

3 The company's submission

The Appraisal Committee (section 9) considered evidence submitted by Lundbeck and a review of this submission by the Evidence Review Group (ERG; section 10).

Clinical effectiveness

3.1 The company conducted a systematic review of the literature to identify studies evaluating the clinical effectiveness and safety of vortioxetine for treating adults having a moderate-to-severe major depressive episode. These adults included those who had not tolerated initial antidepressant treatment or whose condition had responded inadequately to it, and who needed further antidepressant therapy (hereafter referred to as the 'second-line population'). Therefore, the company did not include in its analyses all adults with major depressive disorder as specified in NICE's final scope and vortioxetine's marketing authorisation. It identified 2 phase III randomised controlled trials, REVIVE and TAK318.

3.2 REVIVE was an international (14 European countries including the UK), double-blind, randomised, active-control trial. It included 501 adults with a single episode of moderate-to-severe major depressive disorder or recurrent major depressive disorder whose condition had inadequately responded to monotherapy with a selective serotonin reuptake inhibitor (SSRI) or a serotonin–norepinephrine reuptake inhibitor (SNRI). Patients were randomised 1:1 to flexible doses of vortioxetine (10–20 mg daily; starting dose 10 mg daily), or agomelatine (25–50 mg daily; starting dose 25 mg daily). Patients were assessed weekly during the first 4 weeks of treatment and then every 4 weeks until the end of the

12-week treatment period. A further safety assessment was scheduled 4 weeks after completion or withdrawal from the study. Most patients enrolled into REVIVE were women (74.7%), most were white (99.8%), the mean age was 46.3 years and they had a mean of 2.5 previous major depressive episodes. The company stated that both groups had comparable baseline Montgomery-Åsberg Depression Rating Scale (MADRS) scores and previous antidepressant use. Most patients received the maximum dosage of vortioxetine (20 mg, 64.7%) and agomelatine (50 mg, 71.7%) from weeks 4–12.

3.3 The primary outcome measure in REVIVE was change from baseline in MADRS score at week 8 (MADRS is a rating scale consisting of 10 items, each rated 0 [no symptom] to 6 [severe symptom], contributing to a total score from 0 to 60; the higher the score, the more severe the condition). A 'full analysis set' population (that is, people who were randomised into the study and had a baseline assessment and at least 1 further assessment) was used to analyse the efficacy outcomes. The company tested a primary hypothesis of non-inferiority, and a secondary hypothesis of superiority. Non-inferiority was considered established if the upper bound of the two-sided 95% confidence interval of the difference between treatment groups in MADRS total score at week 8 did not exceed +2 MADRS units compared with agomelatine. The mean change from baseline in MADRS total scores at week 8 were −16.5 and −14.4 points in the vortioxetine group and the agomelatine group respectively. This resulted in a mean difference of −2.16 points in favour of vortioxetine (95% confidence interval [CI] −3.51 to −0.81; see table 1).

3.4 Pre-specified subgroup analyses of the primary outcome were carried out by the company for sex, age, baseline severity, baseline anxiety and class of prior antidepressant. The company stated that

these analyses suggested that vortioxetine improved the MADRS score compared with agomelatine across all pre-specified subgroups.

3.5 The company stated that vortioxetine statistically significantly improved outcomes compared with agomelatine across the analyses of response and relapse outcomes measured by MADRS score (see table 2).

Table 1 Company's analysis of primary outcome in REVIVE

Outcome	Vortioxetine: difference compared with agomelatine			
	Week 8		Week 12	
	MMRM	LOCF, ANCOVA	MMRM	LOCF, ANCOVA
Δ MADRS total score	−2.16*† (−3.51 to -0.81)	−3.1**	−2.03* (−3.45 to -0.60)	−3.5**
Δ=mean change from baseline. † Primary efficacy analysis. *p<0.01; **p<0.001 compared with agomelatine. Vortioxetine: baseline n=252, week 8 n=220, week 12 n=200. Agomelatine: baseline n=241, week 8 n=190, week 12 n=178. Abbreviations: ANCOVA, analysis of covariance; LOCF, last observation carried forward; MADRS, Montgomery-Åsberg Depression Rating Scale; MMRM, mixed model for repeated measures; n, number.				

Table 2 Response and remission in REVIVE

	Response (MADRS)	Remission (MADRS)
Week 8		
Vortioxetine	62%**	41%**
Agomelatine	47%	30%
Adjusted odds ratio (95% CI)	1.81 (1.26 to 2.60)	1.72 (1.17 to 2.52)
Week 12		
Vortioxetine	70%**	55%***
Agomelatine	56%	39%
Adjusted odds ratio (95% CI)	1.83 (1.26 to 2.65)	2.01 (1.39 to 2.90)
*p <0.05; **p <0.01; ***p <0.001 compared with agomelatine. Abbreviations: CI, confidence interval; FAS, full analysis set; LOCF, last observation carried forward; LREG, logistic regression; MADRS, Montgomery-Åsberg Depression Rating Scale.		

3.6 Health-related quality of life was measured at baseline and at weeks 4, 8 and 12 in the REVIVE trial using the EuroQol-5 dimensions survey (EQ-5D, see table 3).

Table 3 EQ-5D summary scores and changes in EQ-5D score from baseline

Assessment	Vortioxetine			Agomelatine			p value
	n	Mean (SD)	Change from baseline*	n	Mean (SD)	Change from baseline*	
Baseline	252	0.53 (0.28)		241	0.55 (0.27)		
Week 4	241	0.70 (0.22)	0.16	233	0.64 (0.27)	0.08	<0.001
Week 8	220	0.76 (0.19)	0.20	189	0.73 (0.23)	0.16	0.03
Week 12	200	0.81 (0.21)	0.25	178	0.78 (0.22)	0.20	0.01

*Based on a mixed model for repeated measures analysis.
Abbreviations: SD, standard deviation; n, number.

3.7 TAK318 was a multicentre (62 centres in USA and Canada), double-blind, randomised, active-control trial including 447 adults with stable major depressive disorder experiencing treatment-emergent sexual dysfunction. Patients were randomised 1:1 to flexible doses of vortioxetine (10–20 mg daily; starting dose 10 mg daily), or escitalopram (10–20 mg daily; starting dose 10 mg daily). Patients were assessed at the end of the 8-week treatment period and had an additional safety assessment 3 weeks after study completion.

3.8 The primary outcome measure in TAK318 was change from baseline in the Changes in Sexual Functioning Questionnaire Short-Form 14 (CSFQ-14) total score after 8 weeks of treatment (total score ranges from 14 to 70; higher scores reflect higher sexual functioning). A 'full analysis set' population was used to

analyse the efficacy outcomes. Sexual functioning improved in both the vortioxetine and escitalopram groups, with a mean difference of 2.2 points in favour of vortioxetine compared with escitalopram (p=0.013).

3.9 There are no head-to-head data comparing vortioxetine with comparators other than agomelatine in the second-line population. Therefore, the company conducted both a Bayesian indirect treatment comparison and a frequentist indirect treatment comparison using the Bucher method for 2 outcomes: rate of remission, and the proportion of people who stop treatment because of adverse events. The company systematically searched the literature and identified the REVIVE trial plus 3 additional multicentre, blinded, randomised, controlled trials comparing: agomelatine with sertraline (Kasper et al., 2010); venlafaxine with citalopram (Lenox-Smith et al., 2008); and bupropion with sertraline or venlafaxine (STAR*D). The company excluded:

- Rosso et al. (2012), which compared bupropion with duloxetine, because it considered the method of randomisation (by day of the week) and blinding (single-blind) inadequate
- 2 placebo-controlled trials because the company's clinical advisers suggested that people who enrol in placebo-controlled trials may be different from those in active-controlled studies, but the company included these trials in a sensitivity analysis.

The company stated that its searches did not identify any evidence that allowed 2 other relevant comparators (fluoxetine or mirtazapine) to be included in the indirect treatment comparison.

3.10 Kasper et al. (2013) was a post-hoc analysis of the 'pre-treated' population from 2 trials of agomelatine in people with major depressive disorder. The number of patients enrolled in each of the

4 trials ranged from fewer than 100 (Kasper) to 789 (STAR*D). The mean age of patients was reported for 3 of the 4 trials and ranged from 41.8 years (STAR*D) to 46.3 years (REVIVE). Baseline severity measured by Hamilton Depression Rating Scale (HAM-D) was between 21 (REVIVE) to more than 31 (Lenox-Smith et al., 2008), but the company considered that the differences between the trials would not have had an impact on the treatment effect. In general, STAR*D enrolled a higher proportion of men who were younger and whose depression was less severe than the populations in the other trials. Outcomes were assessed at different time-points in the trials, from 6 weeks (Kasper) to 14 weeks (STAR*D). Each study measured depressive symptoms (and hence remission) using different scales: MADRS (REVIVE), HAM-D_{17} (Kasper, STAR*D) and HAM-D_{21} (Lenox-Smith). However, the company stated that each trial used clinically accepted cut-off rates for remission, which are generalisable regardless of the scale used.

3.11 The company stated the results of its indirect treatment comparison suggested that vortioxetine works better and is better tolerated than the comparators. The results of the company's indirect treatment comparison are presented in tables 4 and 5. The company stated that it did not assess heterogeneity because of the small number of studies included in the network.

Table 4 Summary of results of company's frequentist indirect treatment comparison

Treatment	Remission rate			People stopping treatment because of adverse events (withdrawal)		
	Rate (%)	Risk difference versus vortioxetine (%)	95% CI	Rate (%)	Risk difference versus vortioxetine (%)	95% CI
Vortioxetine	40.5	–	–	5.9	–	–
Agomelatine	29.5	−11.0	−19.4 to −2.6	9.5	3.6	−1.1 to 8.3
Sertraline	26.1	−14.4	−29.9 to 1.1	18.0	12.1	3.1 to 21.1
Venlafaxine	33.3	−7.2	−24.3 to 9.9	18.2	12.3	0.8 to 23.8
Bupropion	29.8	−10.7	−27.8 to 6.4	24.2	18.3	6.4 to 30.1
Citalopram	23.7	−16.8	−41.1 to 7.5	18.0	12.1	−0.3 to 24.5
Abbreviation: CI, confidence intervals.						

Table 5 Summary of results of company's Bayesian indirect treatment comparison

Treatment	Remission rate			People stopping treatment because of adverse events (withdrawal)		
	Rate (%)	Odds ratio vortioxetine versus comparator (%)	95% CrI	Rate (%)	Odds ratio vortioxetine versus comparator (%)	95% CrI
Vortioxetine	40.5	–	–	5.9	–	–
Agomelatine	29.5	1.63	1.12 to 2.37	9.5	0.60	0.30 to 1.17
Sertraline	25.9	1.95	0.89 to 4.24	29.5	0.15	0.03 to 0.62
Venlafaxine	35.1	1.26	0.51 to 3.07	29.5	0.15	0.03 to 0.65
Bupropion	30.7	1.54	0.62 to 3.77	38.5	0.10	0.02 to 0.46
Citalopram	25.6	1.98	0.59 to 6.60	29.5	0.15	0.02 to 0.86
Abbreviation: CrI, credible intervals.						

3.12 The company presented short-term safety data from REVIVE. About half of patients in each treatment group had 1 or more adverse reaction over the 12-week treatment period. Adverse

reactions with an incidence of 5% or more for vortioxetine or agomelatine respectively were: nausea (16.2% and 9.1%), headache (10.3% and 13.2%), dizziness (7.1% and 11.6%) and somnolence (4.0% and 7.9%). Fewer patients in the vortioxetine group (1.2%) compared with the agomelatine group (1.7%) experienced serious adverse events. Fewer patients stopped treatment because of adverse events in the vortioxetine group (5.9%) than in the agomelatine group (9.5%).

3.13 The company also presented safety data from 5 open-label long-term extension studies including a total of 2587 patients, of which 54% received vortioxetine for 52 weeks or more. The overall incidence of adverse reactions was 74.6%, and was higher in the 15–20 mg dose group (78.9%) than in the 2.5–10 mg group (71.2%).

Cost effectiveness

3.14 The company did not identify any published studies of the cost effectiveness of vortioxetine for treating the second-line population. It submitted a decision tree model with a Markov component to include subsequent treatment switches to third and later lines. It assumed that a patient can be offered 1 of 5 treatments: vortioxetine, agomelatine, citalopram, sertraline and venlafaxine. The company conducted the economic analysis from an NHS and personal social services perspective and chose a time horizon of 12 months so did not discount costs and health effects. A half-cycle correction was applied to the health effects but not the costs in the Markov part of the model (cycle length 2 months).

3.15 The company stated its economic model represented a single major depressive episode. Hypothetical patients entered the model with major depressive disorder that had not responded to initial therapy. The decision tree included:

- an acute phase of treatment of 8 weeks (months 0–2)
- a maintenance phase of 6 months (months 2–8) and
- a recovery phase of 4 months duration (months 8–12).

The time which patients spent in the decision tree varied and depended on whether treatment was successful in each phase. If treatment succeeded in all 3 phases, with remission achieved and sustained to recovery, a hypothetical patient spent the entire 12 months in the decision tree model. The company's economic model also included events in which treatment was not successful (lack of response or adverse events). These events led to a further treatment, that is, to third and subsequent lines of treatment. Patients who did not complete the acute or maintenance phase left the decision tree model and entered the Markov component. In a given cycle of the Markov component, patients could either achieve remission or not. The company assumed that patients remained on treatment for 6 months after they achieved remission in the acute phase unless they experienced a long-term adverse reaction (insomnia, sexual dysfunction or weight gain).

3.16 The company took data on the probability of remission after 8 weeks of treatment (acute phase) from its indirect treatment comparison (see table 4). The company assumed that an person's probability of relapse depended on the line of treatment rather than specific drug: initial second-line treatment (14.2%, from Limosin et al., 2004), third-line treatment (25.0%, from STAR*D), and fourth- plus fifth-line treatment (42.6%, from STAR*D). STAR*D was a prospective, sequentially randomised controlled trial of outpatients with nonpsychotic major depressive disorder who received 1 (n=3671) to 4 (n=123) successive acute treatment steps, including treatment combinations and augmented therapies. Patients who relapsed during the maintenance phase (which the company assumed occurred halfway through this phase) could

switch to third and subsequent lines of treatment. The company assumed that clinicians then assessed these patients for remission 2 months after starting third-line treatment. It took the data reflecting the proportion of patients in remission after each line of treatment from STAR*D: third- (13.7%), fourth- (13.0%) and fifth-line treatment (13.0%). The company considered that patients who had not relapsed after 6 months of maintenance treatment had recovered. These patients stopped treatment and the company assumed that they could not experience recurrent depression.

3.17 Resource use and costs in the company's economic model included those for treatment (drug), adverse events and each health state (that is, monitoring, inpatient and outpatient admissions). The company based drug costs on the list prices from the 'Monthly Index of Medical Specialities'. Dosages in the acute phases were based on the World Health Organization Defined Daily Dose (for example, 10 mg daily for vortioxetine), and dosages in the maintenance phase were based on the mean dose reported at the end of trials included in the company's indirect treatment comparison. The company took data for health state resource use for the acute phases from an unpublished interim analysis of the PERFORM study (n=226, which included people previously untreated) and, for the maintenance phase, from Byford et al. (2011; the General Practice Research Database 2001/06 – 88,935 people with depression and at least 2 antidepressant prescriptions). The company took data for the health state costs from Unit Costs of Health and Social Care (2013) and NHS Reference Costs. The company assumed that no treatments were prescribed to manage adverse events, but that around one-third of people would incur an additional GP visit. Therefore, the company costed all adverse events based on an assumed 0.3 GP visits per patient per adverse event (£13.50).

3.18 To estimate health-related quality of life in the acute phase, the company used EQ-5D data from REVIVE (see table 6). However, for the maintenance phase, the company used EQ-5D data from Sapin et al. (2004). Sapin was a French study that included 250 people with major depressive disorder in primary care, and assessed health-related quality of life at baseline and after 8 weeks of treatment. The company noted that the mean MADRS score at baseline was 32.7 in Sapin compared with 29.1 in REVIVE, which may explain why the baseline EQ-5D score from Sapin was lower than that in REVIVE. The company included disutility values associated with adverse events from Sullivan et al. (2004), and applied them for 3 weeks in the company's base case analysis.

Table 6 Summary of utility values used in company's economic model

Event	Utility value	Comment	Source
Acute phase (0–8 weeks)			
Depression (baseline)	0.54	None	REVIVE
Remission	0.85		
No remission	0.62	Weighted average of people whose depression had not responded to treatment and people whose depression had responded but not remitted at 8 weeks	
Maintenance phase (after 8 weeks)			
Remission	0.85	EQ-5D score for people whose depression had remitted or responded to treatment	Sapin et al., 2004
Relapse/no remission	0.58	EQ-5D score for people whose depression had not responded to treatment	
Recovery	0.85	Assumed equal to remission	
Disutility values (decrements) of adverse events			
Sexual dysfunction	0.049	None	Sullivan et al., 2004
Headache	0.115		
Diarrhoea	0.044		
Somnolence	0.085	Assumed equal to drowsiness	
Nausea	0.065	Assumed average of gastrointestinal adverse events	
Insomnia	0.129	Assumed equal to anxiety	
Dry mouth	0.000	No data available, so company assumed no decrement	Not applicable
Dizziness	0.000		
Sweating	0.000		
Weight gain	0.032	Company calculation	Dixon et al., 2004 & REVIVE

3.19 The company's deterministic cost-effectiveness results for vortioxetine compared with the comparators in the second-line population are presented in table 7.

Table 7 Company's base-case cost-effectiveness results for vortioxetine in people having second-line treatment

Technology	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Venlafaxine	£964	0.675	-	-	-
Vortioxetine	£971	0.694	£7	0.019	£378
Citalopram	£976	0.664	£5	−0.030	Dominated
Sertraline	£977	0.664	£0	−0.001	Dominated
Agomelatine	£1082	0.676	£105	0.012	Dominated

Dominated, fewer QALYs at greater cost than comparator.
Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

3.20 The company explored parameter and structural uncertainty in its economic model by presenting the results of 1-way sensitivity analyses, scenario analyses and a threshold analysis. The 1-way sensitivity analyses suggested the company's cost-effectiveness results were most sensitive to:

- the difference in remission rates at 8 weeks (acute phase) between vortioxetine and each comparator
- GP consultation costs
- the utility value for remission at 8 weeks
- the utility value for relapse after 8 weeks.

However, in all but 2 of the company's 1-way sensitivity analyses, vortioxetine remained dominant or had an incremental cost-effectiveness ratio (ICER) below £15,670 per quality-adjusted life year (QALY) gained. Vortioxetine was dominated by venlafaxine and by citalopram when the lower bound of the 95% confidence interval was included for the differences in remission rates at 8 weeks. The company commented that its scenario analyses showed that its economic model was robust to all of the structural assumptions and remained the most cost-effective treatment. Because the remission rate at 8 weeks was the most influential

driver of the company's cost-effectiveness results, it explored a threshold analysis around this parameter for vortioxetine, see table 8.

Table 8 Company's threshold analysis of remission rate for vortioxetine

Treatment	Remission rate at 8 weeks	£20,000 per QALY gained threshold	Remission rate at 8 weeks	£30,000 per QALY gained threshold
Vortioxetine (base case)	**40.50%**		**40.50%**	
Vortioxetine	30.53%		30.10%	
Venlafaxine	**33.30%**	£20,009	**33.30%**	£29,898
Vortioxetine	27.97%		28.54%	
Agomelatine	**29.50%**	£20,016*	**29.50%**	£29,973*
Vortioxetine	24.53%		24.00%	
Sertraline	**26.10%**	£20,075	**26.10%**	£30,062
Vortioxetine	24.10%		23.55%	
Citalopram	**23.70%**	£20,027	**23.70%**	£29,975

Figures in bold are base case remission rates.

* Threshold ICERs between vortioxetine and agomelatine are based on lower cost and fewer QALYs for vortioxetine, so the ICERs should be interpreted as willingness to accept QALYs lost, not willingness to pay for QALYs gained.

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

ERG critique of clinical effectiveness

3.21 The ERG stated that the reporting of the company's searches were clear and appropriate. The ERG noted that the company presented no evidence to suggest that the relative efficacy between non-SSRIs may vary between first- and second-line use (and beyond). It stated that it would be more appropriate to include the full evidence base for vortioxetine and its comparators, rather than restricting the evidence base from the outset to the second-line population, so excluding 22 of the 24 completed studies of vortioxetine.

3.22 The ERG commented that REVIVE and TAK318 appeared well conducted but raised the following concerns:

- Both trials included comparators of limited relevance to clinical practice in England (NICE has not issued any guidance for agomelatine; NICE technology appraisal 231 [terminated]).
- Both trials were short considering the duration of treatment recommended by NICE to achieve and consolidate remission, so evidence of long-term efficacy was uncertain.
- Both trials evaluated the efficacy of vortioxetine 10–20 mg daily, so the efficacy of the licensed 5 mg daily regimen was uncertain.

3.23 The ERG commented that the population enrolled into REVIVE was broadly representative of the second-line population in England. For example, baseline MADRS scores ranged from 22–43 points, which is consistent with people with moderate-to-severe major depressive disorder. However, the ERG noted that:

- most patients were white (99.8%), which is unlikely to be reflective of the second-line population in England
- 23% of patients had received an SNRI as initial treatment, which is not reflective of clinical practice in England, where SNRI use in first line is negligible
- most patients were recruited from an outpatient psychiatric setting (97.2%)
- the proportion of patients from the UK was small (about 7%).

3.24 The ERG noted that, although the efficacy analyses in REVIVE and TAK318 used a modified intention-to-treat analysis (that is, full analysis set rather than inclusion of all randomised patients), the risk of bias was likely to be low because relatively few patients randomised were excluded.

3.25 The ERG commented that the results from the company's analysis of the primary and secondary outcomes from REVIVE had relatively wide confidence intervals, so the size of the difference in

efficacy between vortioxetine and agomelatine was uncertain (see tables 1 and 2).

3.26 The ERG agreed with the company's assessment of bias for Rosso et al. (2012), so considered it was reasonable to exclude it, but noted it was the only trial that compared vortioxetine with duloxetine. The ERG stated that it was questionable whether Kasper et al. (2013) was suitable for inclusion in the indirect treatment comparison. It stated that it was unclear whether the population consisted entirely of patients receiving second-line treatment, or whether it also included those who had been treated for a previous depressive episode in the last 12 months but were starting first-line treatment for a current major depressive episode.

3.27 The ERG stated that it had significant concerns over the validity of the company's indirect treatment comparison because of the differences in the baseline patient characteristics and severity of depression of the populations across the 4 trials. It also stated that time of outcome assessment between trials (varying from 6–14 weeks) may also affect the results because rates of remission and withdrawal are likely to be time-dependent. The ERG concluded that the heterogeneous nature of data included in the network meant that the results may not be reliable.

3.28 The ERG highlighted that there was little evidence of a statistically significant improvement in the efficacy for vortioxetine compared with the comparators, given that the results from the company's indirect treatment comparison had wide confidence intervals. It stated that the findings in each specific trial drove the results of the company's indirect treatment comparison because of the sparse evidence network (that is, each arm of the network was informed by 1 trial). The ERG noted that basing results on risk differences was potentially inappropriate because they may be sensitive to the

heterogeneity across trials (see table 4). However, it acknowledged that the company's results based on odds ratios were largely consistent (see table 5). The ERG also commented that the results from the company's sensitivity analysis including the 2 placebo-controlled trials were broadly similar to those that excluded them.

3.29 The ERG stated that there was no evidence to suggest the relative efficacy between drugs that were not classified as SSRIs (for example, SNRIs) varied between first- and second-line treatment (and beyond) (see section 3.21). Therefore, it sought further evidence from the company on a first-line population during the clarification stage:

- The ERG re-analysed data from a published meta-analysis of placebo-controlled trials with active reference treatment arms (Pae et al., 2014). Pae compared vortioxetine with agomelatine (1 trial), duloxetine (5 trials) and venlafaxine (1 trial). The ERG noted that both the European Medicines Agency and the company have criticised the use of trials including active references because they are not true randomised comparisons, given that patients whose condition is known to be non-responsive to the reference treatment are excluded, possibly biasing results in favour of the active reference. The ERG accepted the potential for such bias, but did not consider it substantial enough to exclude these trials. The ERG stated that Pae found no evidence of any difference in efficacy between vortioxetine and venlafaxine, and that vortioxetine was less efficacious than duloxetine in reducing depression scores, or achieving response and remission.
- Llorca et al. (2015) published an indirect treatment comparison that included 57 placebo controlled trials of the following drugs: vortioxetine, agomelatine, desvenlafaxine, duloxetine, escitalopram, sertraline, venlafaxine, vilazodone. Llorca found

no evidence of any difference in efficacy between vortioxetine and its comparators. The ERG commented that there was evidence to suggest fewer people stop vortioxetine because of adverse events than other treatments, including sertraline and venlafaxine. The ERG considered that Llorca may represent the most reliable evidence for comparing vortioxetine with other treatments.

3.30 The ERG concluded that, based on all the evidence, vortioxetine is likely to be similar in efficacy to other antidepressants, but may be superior to agomelatine and inferior to duloxetine.

3.31 The ERG agreed that vortioxetine appears generally safe and tolerable in people with major depressive disorder. The ERG stated that, although the incidence of adverse events was high in people taking vortioxetine, most were mild to moderate in nature, and there was no conclusive evidence that they were dose-dependent.

3.32 The ERG also concluded that vortioxetine may have a better overall safety profile than other antidepressants, but sparse comparative data for adverse events prevented the ERG making a firm conclusion.

ERG critique of cost effectiveness

3.33 The ERG stated that the company developed an unnecessarily complicated model structure, and that it was unclear why:

- The company used different modelling approaches in the maintenance and recovery periods, rather than an initial decision tree for the acute phase and then a separate Markov component for all people in the subsequent 10 month period.
- The company assumed different time-points for relapse (after 3 months) and stopping treatment in the maintenance phase because of adverse reactions (after 1 month), which favoured

those treatments with higher acquisition costs. The ERG noted that this introduced inconsistency between the timing of relapse for people within the decision tree and Markov components.

3.34 The ERG commented that basing the decision to change treatments solely on remission data at 8 weeks was an important limitation. It stated that the company's model therefore excluded people whose condition responded to treatment but who had not had full remission and that, in clinical practice, clinicians use response in deciding whether to continue treatments. The ERG commented that the company also used the 8-week remission data to inform decisions to change treatment at 4 weeks in the model. The ERG explained that this ignored the costs of additional treatment for people whose disease responded but did not remit. It also explained that it may have overestimated health benefits for people whose disease remits because it assigned a utility value based on improved health improvements demonstrated over 8 weeks rather than 4 weeks. The ERG concluded that the company's base case may have underestimated vortioxetine's costs and overestimated vortioxetine's benefits.

3.35 The ERG noted:

- The company had assumed that because a person is not at risk of relapse or recurrence in the recovery period, it introduced a potential bias in favour of the most effective initial treatment. The ERG agreed that the risk of relapse may be different in later phases than in earlier phases of the condition, but that assuming no relapse seemed overly optimistic.
- The company had assumed that patients remain on treatment for 6 months after remission in the maintenance phase. The ERG considered that this was reasonable and consistent with NICE's guideline on depression in adults, but was aware that

NICE recommends 2 years of continued treatment in people considered high risk of relapse.

The ERG acknowledged that the company explored both of these assumptions in the company's response to clarification by varying the time-horizon of the model from 8 months (no recovery period) up to 2 years (treatment and monitoring costs continued in the recovery period). The ERG concluded that, although the company's base-case analysis was robust to these scenarios, the ICER for vortioxetine compared with the next most cost-effective treatment was higher than in the company's base-case analysis, suggesting that including these assumptions had potentially favoured vortioxetine.

3.36 The ERG stated a half-cycle correction for both costs and utility values would have been appropriate, rather than for utility values only, because different health states are associated with different costs for consultation or hospitalisation.

3.37 The ERG highlighted that using a 12-month time horizon was reasonable for the 'average' patient because an untreated episode of major depressive order is estimated to last 5–6 months (World Health Organization, 2008). However, the ERG noted that some people may be treated for longer than 12 months and therefore 12 months may not have been sufficient to capture all of the relevant costs and benefits.

3.38 The ERG highlighted that there was uncertainty around whether STAR*D was an appropriate study to inform the prognosis of people with depression whose condition had not remitted after second-line treatment. The ERG considered that STAR*D included treatments that did not reflect the comparators in the model, and that the population of STAR*D was different from the population of REVIVE. It explained that using data from STAR*D for third- and

later lines of treatment imposed a poorer prognosis (that is, lower remission rates and higher relapse rates) than expected for a population with the same characteristics as in REVIVE. The ERG stated that using STAR*D may have made the most effective second-line treatment look even better (that is, vortioxetine in the company's base case analysis).

3.39 The ERG disagreed with the company's decision to use the same utility value for relapse, and people whose condition was not in remission after third or subsequent treatments. This was because they are very different health states. It highlighted that the utility value from Sapin et al (2004), used by the company for people who had not had remission, was lower than the utility value reported for people who had not add remission at week 8 in the REVIVE trial. The ERG considered that it was not necessary to use a different source for the utility values in the maintenance phase, and that using these 2 sources (REVIVE and Sapin et al., 2004) favoured vortioxetine in the company's base-case analysis. It also felt that the relapse health state should have reflected the recurrence of moderate-to-severe major depression and so these people should have returned to their baseline level of utility (that is, 0.54). The ERG proposed alternative utility values for the company's model, see table 9.

Table 9 ERG's preferred utility values

Health state	Company's utility	Company's source	ERG's utility	ERG's source
No remission (0–8 weeks)	0.62	REVIVE	0.67	REVIVE (FAS, MMRM)
No remission (after 8 weeks)	0.58	Sapin (2004)	0.67	
Relapse (after 8 weeks)	0.58	Sapin (2004)	0.54	REVIVE (baseline depression)
Abbreviations: ERG, Evidence Review Group, FAS, full analysis set; MMRM, mixed model for repeated measures.				

3.40 Given the issues highlighted by the ERG around the company's indirect treatment comparison (see sections 3.26 to 3.28), the ERG stated that there was considerable uncertainty associated with the ICERs. It concluded that the company's base-case analysis can only be reliably used for comparisons of vortioxetine with agomelatine.

3.41 The ERG was aware from the World Health Organization (2008) that an untreated major depressive episode lasts on average 5–6 months. The ERG calculated the average duration of a major depressive episode for each treatment included in the company's model based on approximating the mean number of months not spent in the remission and recovery health states. The ERG highlighted that the lowest estimated duration for a major depressive episode for any given treatment in the company's model was for vortioxetine (6.73 months; longer than the 5–6 months stated by the World Health Organization). The ERG explained that this assumed implicitly that people who change treatment have a poorer prognosis compared with the broader major depressive disorder population. This therefore highlighted that the sources used to inform the parameters for remission and relapse for third

and later lines of treatment in the company's model were crucially important (for example, STAR*D).

3.42 The ERG presented deterministic ICERs for several exploratory analyses for second-line treatment that used alternative sources of evidence for the relative effectiveness of vortioxetine compared with its comparators (see section 3.29 and table 10) and used the company's preferred utility values (see table 9).

- Exploratory analysis 1 (see table 11):
  - The dosage of treatment was up-titrated after 8 weeks (maintenance phase).
  - STAR*D was used to inform remission and relapses rates for third- and later lines of treatment.
- Exploratory analysis 2 (see table 12):
  - The same dosage of treatment was used for the acute and mainanteance phases rather than up-tritrated after 8 weeks.
  - STAR*D was used to inform remission and relapses rates for third- and later lines of treatment.
- Exploratory analysis 3 (see table 13):
  - The dosage of treatment was up-titrated after 8 weeks.
  - The remission rate for all treatments used third and subsequent lines of treatment was assumed to be equal to the average of the remisison rates of the second-line comparators. Therefore, the ERG assumed that the absolute rate of remission did not change from third and subsequent lines of treatment.
  - The same rate of relapse was applied for second and subsequent lines of treatment rather than based on the line of treatment (relapse rate taken from Limosin et al., 2004).
- Exploratory analysis 4 (see table 14):
  - The dosage of treatment was up-titrated after 8 weeks

- For third and subsquent lines of treatment, all treatments had the same remission rates. However, the remission rates declined after each line of treatment. The ERG took the average of the remisison rates of the second-line comparators and calculated the remission rates for third and subsequentlines of treatment by applying a proportionate reduction based on the STAR*D trial.
- The same rate of relapse was applied for second and subsequentlines of treatment rather than based on the line of treatment (relapse rate taken from Limosin et al., 2004).

Table 10 ERG's alternative scenarios for relative effectiveness: proportion of remitters at 8 weeks

Treatment	Probability of remission			
	Company submission [from ITC]	ERG scenario 1 [Llorca et al 2014]	ERG scenario 2 [Pae et al 2015]	ERG scenario 3 [equal effectiveness]
Vortioxetine	40.50%	40.50%	40.50%	40.50%
Agomelatine	29.50%	35.81%	26.48%	40.50%
Sertraline	26.10%	–	–	–
Venlafaxine (XR)	33.30%	49.70%	42.52%	40.50%
Duloxetine	–	43.23%	49.30%	40.50%
Citalopram	23.70%	–	–	–
Escitalopram	–	40.74%	–	40.50%

Abbreviations: ERG, Evidence Review Group, ITC, indirect treatment comparison; XR, extended release.

Table 11 ERG exploratory analysis 1 using STAR*D data (with up-titration)

	Costs	QALYs	Incremental		ICER	
			Costs	QALYs	w SSRI	w/o SSRI
					(incremental analyses, in relation to next best)	
ERG scenario 1: Llorca et al. (2015)						
Venlafaxine (XR)	£885	0.736	Ref	ref	ref	ref
Escitalopram	£887	0.729	£3	−0.007	Dominated	–
Vortioxetine	£971	0.733	£83	0.004	Dominated	Dominated
Duloxetine	£1,032	0.730	£61	−0.003	Dominated	Dominated
Agomelatine	£1,069	0.728	£36	−0.002	Dominated	Dominated
ERG scenario 2: Pae et al. (2014)						
Venlafaxine (XR)	£919	0.728	Ref	ref	ref	NA
Vortioxetine	£971	0.733	£52	0.006	£9,191	NA
Duloxetine	£1,017	0.737	£46	0.003	£13,393	NA
Agomelatine	£1,088	0.717	£71	−0.020	Dominated	NA
ERG scenario 3: Equal effectiveness						
Escitalopram	£889	0.729	Ref	ref	ref	–
Venlafaxine (XR)	£929	0.725	£40	−0.003	Dominated	ref
Vortioxetine	£971	0.733	£42	0.008	£18,188	£5,318
Duloxetine	£1,039	0.727	£68	−0.006	Dominated	Dominated
Agomelatine	£1,059	0.734	£20	0.007	£128,927	£128,927

Abbreviations: ICER, incremental cost-effectiveness ratio; NA, not applicable; QALY, quality-adjusted

life year; SSRI, selective serotonin re-uptake inhibitor; XR, extended release; w, with; w/o, without.

Table 12 ERG exploratory analysis 2 using STAR*D data (without up-titration)

	Costs	QALYs	Incremental		ICER	
			Costs	QALYs	w SSRI	w/o SSRI
					(incremental analyses, in relation to next best)	
ERG scenario 1: Llorca et al. (2015)						
Venlafaxine (XR)	£869	0.736	Ref	ref	ref	ref
Escitalopram	£886	0.729	£17	−0.007	Dominated	–
Vortioxetine	£971	0.733	£85	0.004	Dominated	Dominated
Duloxetine	£972	0.730	£1	−0.003	Dominated	Dominated
Agomelatine	£1,026	0.728	£54	−0.002	Dominated	Dominated
ERG scenario 2: Pae et al. (2014)						
Venlafaxine (XR)	£906	0.728	Ref	ref	ref	NA
Duloxetine	£949	0.737	£42	0.009	£4,676	NA
Vortioxetine	£971	0.733	£22	−0.003	Dominated	NA
Agomelatine	£1,057	0.717	£86	−0.017	Dominated	NA
ERG scenario 3: Equal effectiveness						
Escitalopram	£887	0.729	Ref	ref	ref	–
Venlafaxine (XR)	£917	0.725	£29	−0.003	Dominated	ref
Vortioxetine	£971	0.733	£54	0.008	£18,535	£6,899
Duloxetine	£983	0.727	£12	−0.006	Dominated	Dominated
Agomelatine	£1,010	0.734	£28	0.007	£57,955	£57,955

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NA, not applicable; QALY, quality-adjusted life year; SSRI, selective serotonin re-uptake inhibitor; XR, extended release; w, with; w/o, without.

Table 13 ERG exploratory analysis 3 assuming same relapse rate and average remission rate of second-line treatments (with up-titration)

	Costs	QALYs	Incremental		ICER	
			Costs	QALYs	w SSRI	w/o SSRI
					(incremental analyses, in relation to next best)	
ERG scenario 1: Llorca et al. (2015)						
Escitalopram	£706	0.777	Ref	ref	ref	–
Venlafaxine (XR)	£724	0.778	£17	0.001	£15,778	ref
Vortioxetine	£796	0.780	£72	0.002	£36,434	£36,434
Duloxetine	£856	0.777	£60	−0.003	Dominated	Dominated
Agomelatine	£882	0.778	£27	0.001	Dominated	Dominated

ERG scenario 2: Pae et al. (2014)						
Venlafaxine (XR)	£751	0.772	Ref	ref	ref	NA
Vortioxetine	£806	0.778	£55	0.005	£10,394	NA
Duloxetine	£864	0.777	£58	−0.000	Dominated	NA
Agomelatine	£889	0.770	£25	−0.007	Dominated	NA
ERG scenario 3: Equal effectiveness						
Escitalopram	£713	0.775	Ref	ref	ref	–
Venlafaxine (XR)	£752	0.772	£39	−0.003	Dominated	ref
Vortioxetine	£802	0.779	£50	0.006	£27,752	£7,882
Duloxetine	£862	0.774	£60	−0.005	Dominated	Dominated
Agomelatine	£891	0.779	£29	0.005	£196,655	£196,655
Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NA, not applicable; QALY, quality-adjusted life year; SSRI, selective serotonin re-uptake inhibitor; XR, extended release; w, with; w/o, without.						

Table 14 ERG exploratory analysis 4 assuming same relapse rate and average remission rate with second-line use with proportionate reduction based on STAR*D [with up-titration]

	Costs	**QALYs**	**Incremental**		**ICER**	
			Costs	**QALYs**	**w SSRI**	**w/o SSRI**
					(incremental analyses, in relation to next best)	
ERG scenario 1: Llorca et al. (2015)						
Escitalopram	£809	0.751	Ref	ref	ref	–
Venlafaxine (XR)	£813	0.755	£3	0.005	£766	ref
Vortioxetine	£899	0.754	£86	−0.002	Dominated	Dominated
Duloxetine	£955	0.751	£56	−0.002	Dominated	Dominated
Agomelatine	£993	0.750	£38	−0.002	Dominated	Dominated
ERG scenario 2: Pae et al. (2014)						
Venlafaxine (XR)	£848	0.747	Ref	ref	ref	NA
Vortioxetine	£906	0.752	£58	0.004	£13,068	NA
Duloxetine	£951	0.755	£45	0.003	£14,583	NA
Agomelatine	£1011	0.739	£60	−0.016	Dominated	NA
ERG scenario 3: Equal effectiveness						
Escitalopram	£815	0.749	Ref	ref	ref	–
Venlafaxine (XR)	£854	0.746	£39	−0.003	Dominated	ref
Vortioxetine	£904	0.752	£50	0.006	£28,270	£7,992
Duloxetine	£964	0.748	£60	−0.005	Dominated	Dominated
Agomelatine	£993	0.753	£29	0.005	£200,797	£200,797
Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NA, not applicable; QALY, quality-adjusted life year; SSRI, selective serotonin re-uptake inhibitor; XR,						

extended release; w, with; w/o, without.

3.43 Full details of all the evidence are in the Committee papers.

4 Consideration of the evidence

4.1 The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of vortioxetine, having considered evidence on the nature of major depressive disorder and the value placed on the benefits of vortioxetine by people with the condition, those who represent them, and clinical experts. It also took into account the effective use of NHS resources.

4.2 The Committee heard from the clinical and patient experts about the nature of major depressive disorder. The Committee understood from the patient expert that treatment success was measured by a broad range of outcomes including time to remission, reduced incidence of relapse, and improvements in sexual function, sleep quality and cognitive function. The patient expert highlighted that the current options for treating major depressive disorder are associated with different adverse reactions, so having access to a range of treatments was important. The clinical and patient experts commented that major depressive disorder can impair a person's social life and ability to work, and impacts the lives of their families and carers. The patient expert explained that some people may stop treatment early because of a perceived lack of response and adverse reactions, and therefore considered that increasing available information about options would encourage people to seek or continue treatment. The Committee recognised the importance of having a range of treatment options for people with major depressive disorder.

4.3 The Committee discussed the management of major depressive disorder in adults. The Committee understood that major depressive disorder often has a remitting and relapsing course. The Committee heard from the clinical expert that, in general, clinical practice reflects the recommendations in NICE's guideline on depression in adults. These include initial treatment in primary care with a generic selective serotonin reuptake inhibitor (SSRI) such as citalopram and high-intensity psychological support. NICE's guideline goes on to recommend that if a person's major depressive episode does not adequately respond, or if the person does not tolerate first-line treatment, clinicians and patients should consider a different SSRI or a better-tolerated, newer-generation antidepressant. The clinical expert stated that most people in the NHS would receive escitalopram (also an SSRI) second line, but treatment choice was influenced by treatment history (for example, number of previous therapies, first or recurrent episode of depression) and presence of specific signs and symptoms. The clinical expert further explained that in clinical practice, people with:

- suspected bipolar disorder may receive fluoxetine (however, the Committee was aware that the company considered only unipolar depression in its submission)
- low energy levels may receive venlafaxine (the Committee was aware that the company stated that venlafaxine is the most commonly used serotonin-norepinephrine reuptake inhibitor [SNRI] at second line)
- agitation may receive mirtazapine because of its sedative effect (but mirtrazapine is associated with weight gain so people may instead receive agomelatine).

The Committee was aware that the NICE's guideline depression in adults gave general practitioners the option to prescribe second-line treatments in primary care (for example, escitalopram or an

SNRI). The Committee further heard from the clinical expert that people with difficult-to-treat, severe depression who needed second- or third-line treatment with an antidepressant from another pharmacological class would be referred to secondary care (for example, psychiatric outpatient clinics).

4.4 The Committee considered the likely position of vortioxetine in the treatment pathway. The Committee noted that vortioxetine has a marketing authorisation in the UK for treating 'major depressive episodes in adults'. However, it noted that the company had not submitted clinical and cost-effectiveness evidence for this population, but only for people with moderate-to-severe major depressive disorder whose condition had responded inadequately in terms of efficacy or tolerability to first-line treatment, and who needed second-line treatment. The Committee heard from the clinical expert that vortioxetine would not be used first line, but was likely to be used second line (in the treatment pathway in the position proposed by the company) and also third line. The clinical expert explained that this was because vortioxetine's tolerability and efficacy are comparable with other antidepressants categorised in NICE's guideline on depression in adults as 'better-tolerated newer generation antidepressants'. The clinical expert stated that vortioxetine was more likely to be prescribed in secondary care than in primary care because its price is higher than other antidepressants. The Committee understood that clinicians would like to use vortioxetine in the secondary care setting for people whose major depressive episode was likely to benefit from second- or third-line treatment (that is, after SSRI therapy) with a 'newer-generation, better tolerated antidepressant'.

Clinical effectiveness

4.5 The Committee reviewed the clinical trial evidence submitted by the company, and agreed that the REVIVE trial comparing vortioxetine with agomelatine was of good quality. However, it noted that a key issue highlighted by the Evidence Review Group (ERG) was the generalisability of the results from REVIVE to people diagnosed with a major depressive episode whose episode had responded inadequately to a course of SSRI antidepressants, that is, the second-line population on which the company focused its evidence submission. The Committee heard from the clinical expert that agomelatine was a reasonable comparator for vortioxetine in a trial setting because it is not sedative. The Committee understood that agomelatine is not widely used in clinical practice in the NHS, but is used as an alternative treatment for some people for whom mirtazapine is not appropriate because it is associated with weight gain. The Committee agreed that, because of agomelatine's limited use, the comparison of vortioxetine with agomelatine was of limited relevance to clinical practice in England. The Committee considered whether the previous treatments received by the REVIVE population were generalisable to clinical practice in England. The Committee was also aware that over 20% of patients in REVIVE received initial treatment with an SNRI rather than an SSRI as recommend by NICE's guideline on depression in adults, and agreed that this did not reflect clinical practice in England. The Committee noted that the number of people recruited to the REVIVE trial from the UK was small (approximately 7%), and agreed that the variation in managing major depressive disorder across countries may limit the applicability of the trial results to patients in England. The Committee concluded that the results from the REVIVE trial were not generalisable to most patients in routine clinical practice in England.

4.6 The Committee considered the results from the REVIVE trial. The Committee heard from the company that it used a 'full analysis set' rather than an intention-to-treat analysis to assess the outcomes, in accordance with the Committee for Medicinal Products for Human Use guidelines for non-inferiority trials. The Committee commented that it preferred to see outcomes analysed using an intention-to-treat analysis but it was aware that few patients were excluded from the 'full analysis set' analysis in the REVIVE trial. The Committee noted that the primary outcome in REVIVE was the change in severity of depressive symptoms measured by the mean change from baseline in the Montgomery-Åsberg Depression Rating Scale (MADRS) at 8 weeks. The mean MADRS score was 2.16 points lower with vortioxetine compared with agomelatine. The Committee also noted that vortioxetine showed a statistically significant improvement in both response and remission rates (secondary outcomes) compared with agomelatine. The Committee discussed what size of changes in depressive symptom severity scores clinicians and patients consider clinically important. The Committee heard from the clinical expert that the mean change from baseline in total MADRS score was not a useful outcome measure for judging whether a clinically important difference was observed because the MADRS included 10 items for measuring depressive symptoms. The clinical expert explained that a reduction in 1 item of the MADRS by 2 or more points would generally be considered clinically meaningful. The Committee agreed that achieving remission and avoiding relapse were much more useful outcomes than the mean change in a person's depressive symptom severity score for measuring success of treatment in clinical practice.

4.7 The Committee discussed the company's indirect treatment comparison in the second-line population. The Committee was

concerned that only 4 trials comprised the evidence network, that is only 1 for each treatment comparison. The Committee was also aware that 1 of these trials (Kasper et al., 2013) included people who may not have been changing to another treatment for a major depressive episode but starting first-line treatment for a recurrent major depressive episode. The company acknowledged that the population included in Kasper may not be comparable with the other populations included in the evidence network, or consistent with the population specified in its decision problem. The Committee considered that the patient populations between the trials differed in baseline severity of depression. The Committee was aware that the company's indirect treatment comparison reported remission rates and the proportion of people stopping treatment because of adverse events, both of which depend on trial duration, which differed between the trials included in the network. The Committee concluded that, because of the evidence base, the company's indirect treatment comparison was not sufficiently robust for estimating the clinical effectiveness of vortioxetine compared with other antidepressants for people having second-line treatment.

4.8 The Committee discussed whether evidence from the first-line treatment population was relevant for informing the relative effectiveness of vortioxetine compared with other antidepressants for people having second-line treatment, as positioned by the company. The Committee heard from the company that, although there is a paucity of evidence for vortioxetine used second line, the company chose not to use data from its trials, including first-line treatment, because it claimed that the relativeness effectiveness changes across lines of treatment. The Committee was aware that the ERG considered that the company did not provide sufficient evidence that the relative effectiveness differs between non-SSRIs

within each line of treatment, but the ERG accepted that the absolute effectiveness may change between each line of treatment. The Committee heard from the clinical expert that, in clinical practice, the absolute effectiveness of each antidepressant is likely to decline with each subsequent line of treatment. It heard this is because there will always be some people whose major depressive episode is difficult-to-treat (that is, treatment-resistant) and therefore unlikely to remit or respond. However, the clinical expert noted that the relative effectiveness of the antidepressants compared with one another may also change at each subsequent line of treatment. The clinical expert explained that depression which does not respond to 1 or 2 SSRIs may be mediated by different receptors, so the relative effectiveness of treatments with a different mechanism may differ across subsequent lines of treatment. The ERG acknowledged that the relative effectiveness may reduce in clinical practice at second or later lines of treatment compared with first-line treatment, particularly for SSRIs when compared with antidepressants of a different class. However, it emphasised that there was no evidence available to support a declining relative effect of treatment between drugs other than SSRIs. The Committee was also aware that NICE's guideline on depression in adults concluded that "the evidence for the relative advantage of switching either within or between classes is weak" and "that evidence from primary efficacy studies of existing treatments should also be considered" when making decisions about second and subsequent lines of treatment. On balance, the Committee concluded that evidence from trials in the first-line population was relevant to informing the relative effectiveness of vortioxetine compared with other antidepressants for second and subsequent lines of treatment.

4.9 The Committee discussed alternative sources (Pae et al., 2015 and Llorca et al., 2014) presented to it by the company to estimate the relative effectiveness of vortioxetine compared with other antidepressants. The Committee was aware that these meta-analyses only included a population being treated first line. The Committee noted that the absolute remission rates for vortioxetine were lower than for some of the other antidepressants included in Pae and Llorca (see table 10). It noted that this was not consistent with the company's indirect treatment comparison, which estimated vortioxetine to be the most effective treatment option (see table 10). The Committee appreciated that the 2 studies took different methodological approaches (see section 3.29). It heard from the ERG that each analysis was subject to a number of biases (for example, Pae included trials with active reference arms), but that the ERG considered Llorca to be the most credible. The Committee was aware that Llorca included more treatment options and trial evidence than Pae, and also used indirect evidence to inform the estimates of relative effectiveness (rather than only direct evidence as carried out by Pae). The Committee concluded that the estimates of relative effectiveness in each analysis were subject to uncertainty but, of the available sources, Llorca had the fewest weaknesses for informing the relative effectiveness of vortioxetine compared with other antidepressants.

4.10 The Committee discussed the relative effectiveness evidence available for vortioxetine compared with other antidepressants. The Committee acknowledged that the available evidence was limited, and that none of the analyses it had seen (that is, the company's indirect treatment comparison, Pae et al. 2015, Llorca et al. 2014), estimated statistically significant differences between vortioxetine and the other antidepressants for achieving remission (other than compared with agomelatine, a comparator not widely used in the

NHS). The Committee highlighted that some differences between the absolute rates of remission estimated in each source of evidence could be considered clinically significant despite the lack of statistical significance (likely to be driven by the small trial populations and sparse nature of the available evidence base). The Committee was aware that Pae (not sponsored by the company) concluded that vortioxetine was “more effective than placebo but the difference was of doubtful clinical significance”, and that Llorca (sponsored by company) concluded that vortioxetine was "comparable or favourable" in efficacy and tolerability compared with other antidepressants. Furthermore, the Committee noted that the evidence for vortioxetine in people having second-line treatment included trials only of short duration, so the treatment effect of vortioxetine after 8 weeks was uncertain. The Committee concluded that no convincing evidence existed to show that vortioxetine was any more or less effective than other antidepressants.

4.11 The Committee discussed the adverse effects associated with vortioxetine and the other antidepressants. The Committee noted that the company’s indirect treatment comparison, Pae et al. (2015), and Llorca et al. (2014) measured the odds of stopping treatment because of adverse events. The Committee was aware that some patients may stop treatment for reasons other than adverse events, and that some patients tolerate adverse events but do not stop treatment. The Committee noted that the company had submitted safety data for vortioxetine in patients treated first line and second line. The Committee agreed that safety data are transferable across lines of treatment. The Committee understood that the dose of vortioxetine generally increases over time and that the long-term safety data suggested that the overall incidence of adverse reactions was higher in people taking 15–20 mg of

vortioxetine daily compared with 5 mg of vortioxetine daily. The Committee noted that the company did not present any evidence comparing the adverse effects associated with vortioxetine with those associated with other antidepressants in the broader population of people with major depressive disorder. The Committee was aware that the TAK318 trial, which the company did not include in its indirect comparison or modelling, showed that vortioxetine improved sexual function in people with sexual dysfunction more than escitalopram. The Committee agreed that the long term adverse effect profile of vortioxetine compared with commonly used antidepressants in England was uncertain. However, it accepted that the available evidence suggested vortioxetine leads to a lower probability of stopping treatment and fewer adverse effects than most other antidepressants in the short term. The Committee concluded that, based on the available evidence, albeit sparse, vortioxetine may have a better overall safety profile than other antidepressants.

Cost effectiveness

4.12 The Committee discussed the company's economic model and cost-effectiveness results. The Committee was aware that the company submitted cost-effectiveness results only for vortioxetine as a second-line treatment, and therefore could not make a recommendation for vortioxetine for treating all people included in the marketing authorisation. The Committee was concerned that the company's model structure lacked face validity and therefore made assessing the cost effectiveness of vortioxetine compared with other antidepressants difficult. It was concerned about several structural uncertainties identified during the ERG's critique, notably that the company:

- incorporated an overly complicated structure that used 2 different modelling approaches for the maintenance and recovery phases, which introduced inconsistencies between the timings of particular events, for example, time to relapse
- used a time horizon of 12 months that was not sufficient to reflect duration of treatment, or for capturing all costs and benefits
- defined treatment success, or decisions to change treatment, solely on whether a person's condition remits or not (see section 4.13)
- used remission data from the trials after 8 weeks of second-line treatment to inform decisions to change to another treatment after 4 weeks in the model (see section 4.13)
- underestimated the duration of maintenance therapy by excluding people with a severe or recurrent major depressive episode (see section 4.13)
- assumed that all people having second-line treatment for a major depressive episode would be treated in primary care (see section 4.13) and none in secondary care
- assumed that people did not experience a relapse of their condition during the recovery phase, which seemed overly optimistic
- did not appropriately model the rates of remission and relapse for third and subsequent lines of treatment (see section 4.16)
- applied a half-cycle correction only to utility values, and not to costs, in the Markov component (see section 4.16).

Because the company's model restricted the decision problem and clinical pathway to people having second-line treatment for a major depressive episode the Committee was also unable to judge with any certainty whether vortioxetine could be considered as a cost-

effective use of NHS resources for third and subsequent lines of treatment.

4.13 The Committee discussed the costs and resource-use values included in the company's economic model. The Committee noted that the dose of second-line treatment was up-titrated after 8 weeks in the company's economic model, and was aware that this may reflect clinical practice in people who tolerate, and whose depression responds to, treatment. Moreover, the Committee understood from the ERG's exploratory analysis that assuming an alternative scenario in which the dose of second-line treatment did not increase after 8 weeks had little impact on the incremental cost-effectiveness results. The Committee was aware that continuing treatment in the company's model was based on whether a person's depression remits or not. The Committee heard from the clinical expert that people with major depressive disorder whose condition responds, but does not remit, after 8 to 10 weeks of treatment would be treated for a further 4 weeks with augmentation therapy in clinical practice (see sections 1.8.1.5 to 1.8.1.9 of NICE's guideline on depression in adults). The Committee also noted that the company modelled remission data from the trials after 8 weeks of second-line treatment to inform decisions to change treatment after 4 weeks of second-line treatment in the company's model. It noted that this underestimated the treatment cost and overestimated the health benefits of people whose depression remits. The Committee was also aware that the company's model assumed 6 months of maintenance therapy because the company considered this to be in line with the recommendations in NICE's guideline depression in adults. However, the Committee was aware that this guideline recommends treatment for up to 2 years in people at high risk of relapse. It heard from the clinical expert that, in England about 30–50% of people experiencing their first major

depressive episode would stop treatment after 6 months, but that people experiencing a recurrent major depressive episode would receive treatment for up to 2 years. The Committee acknowledged that the company had provided a scenario that extended the time horizon of its model to 24 months and assumed treatment continued for up to 2 years. However, the Committee noted that this scenario showed that the incremental costs of vortioxetine increased at a higher rate relative to the other antidepressants and was flawed by structural uncertainties. The Committee agreed that assumptions about when to continue or change treatment in the company's model did not reflect clinical practice in England. The Committee inferred that, had the company modelled more realistic assumptions, then the costs associated with each treatment strategy would likely have increased and, because the list price of vortioxetine is higher than most other antidepressants, this would have disproportionately disadvantaged vortioxetine. The Committee was also aware that the company's model assumed that all people with major depressive disorder remain in primary care after first-line treatment. It noted that, in clinical practice in England, this is not consistent with where vortioxetine is likely to be given (that is, secondary care; see section 4.4). The Committee concluded that the values for cost and resource use included in the company's model did not reflect the pathway of care for people for whom vortioxetine would be considered appropriate, and also excluded people with recurrent major depressive disorder who need intensive treatment (for example, maintenance therapy for up to 2 years).

4.14 The Committee was aware that the company did not assume that treating the population included in its submission lowered the risk of suicide, so any modelled gains in quality-adjusted life years (QALYs) only reflected a difference in utility (that is, health-related

quality of life). The Committee discussed whether it was appropriate for the company to use 2 separate sources of evidence for the utility values in its economic model. The Committee was aware from the ERG that the utility value from Sapin et al. (2004) used in the company's model for people whose condition does not remit was lower than the utility value for people whose condition did not remit by 8 weeks in the REVIVE trial. Furthermore, the Committee understood that the utility value from Sapin chosen by the company for people whose depression subsequently relapsed was higher than the baseline utility value of patients in the REVIVE trial. The Committee agreed that the company's chosen utility values for 'non-remission' and 'relapse' were not internally consistent or consistent with the company's model structure. The Committee concluded that it preferred the company to use the EQ-5D data, collected as a secondary outcome measure in REVIVE, because it represented the best evidence available and more closely reflected the population included in the company's model.

4.15 The Committee discussed the company's approach to modelling adverse events. The Committee was aware that the company based adverse event rates for vortioxetine and its comparators on absolute rates reported from individual trials. The Committee noted that it accepted that safety data would be transferable across lines of treatment (see section 4.11). However, it was uncertain whether the company's approach to modelling adverse events was appropriate, given differences in the baseline severity of depression in the trials' populations for vortioxetine and its comparators. On balance, the Committee recognised that vortioxetine was likely to lead to fewer adverse events compared with other antidepressants. The Committee also noted that the company assigned no decrease in health-related quality of life because of several adverse events

(for example, dry mouth, dizziness), and agreed that this was unlikely to reflect reality. However, the Committee was aware that, for these adverse events, the incidence rates were generally lower for vortioxetine than the other antidepressants, so the company's approach underestimated the benefits of vortioxetine. The Committee also noted that the company had not considered adverse events in people receiving third and subsequent lines of treatment. The Committee agreed that, because a substantial proportion of people receive therapy after second-line use in the company's model, this led to further uncertainty around the cost-effectiveness results. The Committee concluded that it would have preferred the company to justify its approach for modelling adverse reactions, but appreciated that data for antidepressants in comparable populations were likely to be sparse.

4.16 The Committee discussed the company's approach to modelling remission and relapse rates for people having third-line and later lines of treatment other than vortioxetine. The Committee heard from the clinical expert that there was limited evidence available to inform the prognosis for people having third and subsequent lines of treatment (that is, rates of remission or relapse), and that the STAR*D trial provided the best available data. The Committee accepted this, but understood that STAR*D included treatments that did not reflect treatments commonly used in England, and that the population was different from the population in REVIVE. The Committee also appreciated that the effectiveness of third and subsequent lines of treatment was independent of the initial second-line treatment strategy, but the proportion of patients that subsequently switched to third line treatment differed depending upon the effectiveness of the initial second-line treatment. It agreed that the company's approach to using the STAR*D data lead to considerable uncertainty. The Committee noted that the company

should have been more critical of the data in its evidence submission and explored alternative scenarios around the assumptions given STAR*D's influence on the cost-effectiveness results. The Committee acknowledged that the chance of achieving remission would decrease with each line of treatment given and that the population that needs many lines of treatment would include people who were more likely to have difficult-to-treat and treatment-resistant depression. It noted that the company used the absolute rates of remission from the STAR*D trial, which included a population that differed from the REVIVE trial. However, the Committee would have been preferred the company to have applied a proportionate reduction in the rates of remission used for third and subsequent lines of treatment, as seen in the STAR*D trial, to the remission data used for second-line treatment, as explored in the ERG's exploratory analysis 4 (see section 3.42). The Committee noted that company's assumed that the rate of relapse did not differ between second-line treatment but did differ between third and subsequent lines of treatment. The Committee would have preferred the company to have assumed that the rate of relapse was independent of treatment line as considered more appropriate by the ERG. The Committee acknowledged the views of the clinical expert that most people treated with vortioxetine would be seen in the secondary care setting (having been referred for second or third-line treatment), and people whose condition did not remit but had responded to treatment would continue treatment for a further 4 weeks (that is, for a total of at least 12 weeks on second-line treatment). The Committee therefore considered that the company's model may have overestimated the average number of therapies a person received over the time horizon of the model. The Committee also noted that the company's Markov model for subsequent treatment did not apply the half-cycle correction to costs but only to utility values. The Committee highlighted that the

company's approach was not appropriate because most changes in health-related quality of life are a result of changes in health states, which also carry with them different consultation or hospitalisation costs. It therefore agreed that the company's approach for using a half-cycle correction should have been consistent for costs and utility values. The Committee concluded that the company's approach to modelling remission and relapse rates for people having third-line and later lines of treatment was not appropriate.

4.17 The Committee discussed the cost-effectiveness results presented by the company and the ERG's exploratory analyses. The Committee noted that the company's base-case results were not sensitive to changes to most parameters. It also noted that, in all but 2 of the company's scenario analyses, vortioxetine dominated (more effective and less costly) or had an incremental cost-effectiveness ratio (ICER) below £16,000 per QALY gained compared with other antidepressants (see section 3.20). It was aware that, in the company's model, vortioxetine was estimated to be the most effective treatment strategy and that this was based on the company's indirect treatment comparison (see section 4.7). The Committee was aware that, when vortioxetine was not considered the most effective treatment strategy, as estimated by Llorca et al. (2014) and Pae et al. (2015), or specifically when vortioxetine was assumed to be as effective as other antidepressants, the ICERs for vortioxetine were shown to be extremely unstable because of the small differences in incremental QALYs (that is, highly sensitive to the parameters used for the rates of remission and relapse). The Committee understood from the ERG that the small differences between the incremental QALYs were driven partly by the short time horizon of the model and partly by the data suggesting that vortioxetine was not any more or less effective than other antidepressants (but reflecting the small observed differences in

absolute effects). The ERG explained that, given that there were no substantial differences in depressive severity symptoms scores between the antidepressants reported in the trials included in the company's indirect treatment comparison (Pae or Llorca), it was not surprising that the incremental QALYs were equally small. The Committee concluded that it needed to take into account the instability of the ICERs in its decision-making.

4.18 The Committee noted that the company's ICERs and the ERG's exploratory analyses were estimated from deterministic analyses rather than from probabilistic analyses. The Committee acknowledged the length of time needed to run probabilistic analyses was longer than for deterministic analyses. However, it stated that the company had not presented the ICERs for deterministic and probabilistic analyses and therefore it was unable to assess whether the company's economic model was linear or non-linear in nature. The Committee concluded that it preferred probabilistic ICERs presented within a fully incremental analysis and as pairwise comparisons, as defined in NICE's Guide to the methods of technology appraisal (2013).

4.19 The Committee discussed whether it could recommend vortioxetine as a second-line treatment option for treating major depressive episodes. The Committee acknowledged that the company had used the best available evidence to model subsequent treatment and used EQ-5D utility data as preferred by NICE in its Guide to the methods of technology appraisal (2013). However, the Committee noted that it preferred the way in which the ERG applied these data in the ERG's exploratory analyses (scenario 4). The Committee highlighted that in its preferred analysis, vortioxetine was dominated by venlafaxine, and that across all scenarios, there was considerable uncertainty associated with the ICERs because of the company's economic model structure. Furthermore, the

Committee emphasised that there was no convincing clinical-effectiveness evidence to show that vortioxetine was any more or less effective than other antidepressants (see section 4.10), but vortioxetine had a higher acquisition cost than other antidepressants for second-line treatment of major depressive episodes such as venlafaxine, escitalopram and sertraline. The Committee highlighted the unstable nature of the ICERs (see section 4.17) and its view that, if the structural uncertainties were addressed, the incremental costs for vortioxetine would increase disproportionately relative to other antidepressants given its higher acquisition cost (see section 4.13). The Committee concluded that it could not recommend vortioxetine as a cost-effective use of NHS resources for second-line treatment for major depressive episodes.

4.20 The Committee acknowledged that having more treatment options was important for patients, and that vortioxetine may be beneficial for some people. However, the Committee agreed that the company had not sufficiently explored how vortioxetine could be used in the NHS, and its economic model did not reflect how antidepressants are used in clinical practice. The Committee noted the views of the clinical and patient experts, and recognised that vortioxetine may be a valuable treatment option after second-line treatment, or as an option in a secondary care setting, or for people at high risk of adverse events. However, it commented that the clinical and cost-effectiveness evidence submitted by the company prevented it from making a recommendation for any other population than people who needed second-line treatment. The Committee was minded not to recommend vortioxetine for treating adults with major depressive episodes as a cost-effective use of NHS resources. The Committee requested the following further clarification and analyses from the company to address the issues identified:

- a cost-effectiveness analysis of vortioxetine compared with relevant treatment options third line and beyond in the treatment pathway for major depressive disorder, for example, after selective serotonin reuptake inhibitors and 1 serotonin–norepinephrine reuptake inhibitor (venlafaxine), or as an option in a secondary care setting, or in patients who have had multiple adverse reactions
- consideration of these subgroups and to include probabilistic analyses that:
  - incorporate the broader evidence base for antidepressants, including at first-line treatment
  - define treatment success, and decisions to switch treatment, by remission and response
  - use the time point in which patients change to another treatment from the trials for the time point in the model (for example, 8 weeks rather than 4 weeks)
  - consider that people may receive treatment for up to 2 years (for example, to consolidate response)
  - include a risk of relapse at all stages of depression
  - use utility values from REVIVE
  - include a 24-month time horizon
  - present pairwise comparisons and incremental analyses for the probabilistic cost-effectiveness estimates
- disaggregated results for each of the pairwise comparisons.

4.21 The Committee discussed whether vortioxetine could be considered innovative, and whether the company's economic analysis had captured all changes in health-related quality of life. The Committee noted that the company considered vortioxetine innovative because: it reduces cognitive dysfunction independent of its effect on MADRS; it minimises impact on social relationships; it reduces symptoms associated with stopping treatment; and it

provides benefits related to health-related quality of life underestimated by the EQ-5D instrument. The Committee acknowledged that vortioxetine may be a valuable treatment option for people with a major depressive order experiencing cognitive dysfunction. However, it noted that the EQ-5D data from REVIVE reported for the vortioxetine and agomelatine groups did not suggest that the average utility was notably different between treatments. The Committee also acknowledged that, in general, the benefits of mental health conditions relative to other conditions may be underestimated by the EQ-5D instrument. However, the Committee considered that any shortcomings in the EQ-5D would impact each treatment option included in the company's economic analysis similarly, particularly because there was no convincing evidence to suggest that vortioxetine was any more or less effective than its comparators. The Committee concluded that these benefits were sufficiently captured within the company's economic modelling.

Summary of Appraisal Committee's key conclusions

TAXXX	**Appraisal title:**	**Section**
Key conclusion		
The Committee is minded not to recommend vortioxetine within its marketing authorisation, that is, for treating major depressive episodes in adults. The Committee recommends that NICE requests further clarification and analyses from the company, which should be made available for the second Appraisal Committee meeting.		1.1, 1.2, 4.20
Current practice		

Clinical need of patients, including the availability of alternative treatments	The Committee recognised the importance of having a range of treatment options for people with major depressive disorder. The clinical expert stated that most people in the NHS would receive escitalopram (also an SSRI) second line, but treatment choice was influenced by treatment history (for example, number of previous therapies, first or recurrent episode of depression) and presence of specific signs and symptoms.	4.2 4.3
The technology		
Proposed benefits of the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	The Committee noted that the company considered vortioxetine innovative because: it reduces cognitive dysfunction independent of its effect on MADRS; it minimises impact on social relationships; and it reduces symptoms associated with stopping treatment.	4.21
What is the position of the treatment in the pathway of care for the condition?	The Committee understood that clinicians would like to use vortioxetine in the secondary care setting for people whose major depressive episode was likely to benefit from second- or third-line treatment with a 'newer-generation, better tolerated antidepressant'.	4.4
Adverse reactions	The Committee concluded that, based on the available evidence, albeit sparse, vortioxetine may have a better overall safety profile than other antidepressants.	4.11

Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The Committee agreed that the REVIVE trial comparing vortioxetine with agomelatine was of good quality. The Committee concluded that, because of the number and nature of trials included, the company's indirect treatment comparison was not sufficiently robust for estimating the clinical effectiveness of vortioxetine compared with other antidepressants for people having second-line treatment. On balance, the Committee concluded that evidence from trials in the first-line population was relevant to informing the relative effectiveness of vortioxetine compared with other antidepressants for second and subsequent lines of treatment. The Committee concluded that the estimates of relative effectiveness in each analysis were subject to uncertainty but, of the available sources, Llorca had the fewest weaknesses. The Committee noted that the evidence for vortioxetine for second-line treatment included trials only of short duration, so the treatment effect of vortioxetine after 8 weeks was uncertain.	4.5 4.7 4.8 4.9 4.10
Relevance to general clinical practice in the NHS	The Committee concluded that the results from the REVIVE trial were not generalisable to most patients in routine clinical practice in England.	4.5

Uncertainties generated by the evidence	The Committee was concerned that only 4 trials comprised the evidence network, that is, only 1 for each treatment comparison. The Committee considered that the patient populations between the trials differed in baseline severity of depression. The Committee was aware that the ERG considered that the company did not provide sufficient evidence that the relative effectiveness differs between non-SSRIs within each line of treatment, but the ERG accepted that the absolute effectiveness may change between each line of treatment. The Committee heard from the ERG that Pae and Llorca comparison were subject to a number of biases, but that the ERG considered Llorca to be the most credible.	4.7 4.8 4.9
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	Not applicable	-
Estimate of the size of the clinical effectiveness including strength of supporting evidence	The Committee concluded that no convincing evidence existed to show that vortioxetine was any more or less effective than other antidepressants.	4.10
Evidence for cost effectiveness		

Availability and nature of evidence	The Committee was aware that the company submitted cost-effectiveness results only for vortioxetine as a second-line treatment. The Committee commented that the clinical and cost-effectiveness evidence submitted by the company prevented it from making a recommendation for any other population than people who needed second-line treatment.	4.12 4.20
Uncertainties around and plausibility of assumptions and inputs in the economic model	The Committee was concerned that the company's model structure lacked face validity and therefore made assessing the cost effectiveness of vortioxetine compared with other antidepressants difficult. The Committee concluded that the values for cost and resource use included in the company's model did not reflect the pathway of care for people for whom vortioxetine would be considered appropriate, and also excluded people with recurrent major depressive disorder who need intensive treatment. The Committee concluded that the company's approach to modelling remission and relapse rates for people having third-line and subsequent lines of treatment was not appropriate.	4.12 4.13 4.16

Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The Committee concluded that it preferred the EQ-5D data collected in REVIVE, because it represented the best evidence available and more closely reflected the population included in the company's model. The Committee concluded that vortioxetine's benefits were sufficiently captured within the company's economic modelling.	4.14 4.21
Are there specific groups of people for whom the technology is particularly cost effective?	Not applicable	-

What are the key drivers of cost effectiveness?	The Committee inferred that, had the company modelled more realistic assumptions, then the costs associated with each treatment strategy would likely have increased and, because the list price of vortioxetine is higher than most other antidepressants, this would have disproportionately disadvantaged vortioxetine. The Committee was aware that, when vortioxetine was not considered the most effective treatment strategy, as estimated by Llorca and Pae, or specifically when vortioxetine was assumed to be as effective as other antidepressants, the ICERs for vortioxetine were shown to be extremely unstable because of the small differences in incremental QALYs (that is, highly sensitive to the parameters used for the rates of remission and relapse).	4.13 4.18
Most likely cost-effectiveness estimate (given as an ICER)	The Committee highlighted that in its preferred analysis, vortioxetine was dominated by venlafaxine, and that across all scenarios, there was considerable uncertainty associated with the ICERs because of the company's economic model structure.	4.19
Additional factors taken into account		
Patient access schemes (PPRS)	Not applicable	-
End-of-life considerations	Not applicable	-

Equalities considerations and social value judgements	Potential equality issues raised during the appraisal could not be addressed through NICE technology appraisal guidance.	-

5 Implementation

5.1 Section 7(6) of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within [insert number] months of its date of publication. The normal period of compliance, of 3 months, has been extended for this technology because [insert reason]. This extension is made under Section 7(5) of the Regulations.

5.2 NICE has developed tools [link to www.nice.org.uk/guidance/TAXXX] to help organisations put this guidance into practice (listed below). [NICE to amend list as needed at time of publication]

- Slides highlighting key messages for local discussion.
- Costing template and report to estimate the national and local savings and costs associated with implementation.
- Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
- A costing statement explaining the resource impact of this guidance.
- Audit support for monitoring local practice.

6 Related NICE guidance

Details are correct at the time of consultation and will be removed when the final guidance is published. Further information is available on the NICE website.

Published

- Agomelatine for the treatment of major depressive episodes (terminated appraisal). NICE technology appraisal 231 (2011).
- Depression in adults quality standard. NICE quality standard 8 (2011).
- Depression in adults. NICE clinical guideline 90 (2009).
- Depression in adults with a chronic physical health problem. NICE clinical guideline 91 (2009).
- Computerised cognitive behaviour therapy for depression and anxiety. NICE technology appraisal guidance 97 (2006).

7 Proposed date for review of guidance

7.1 NICE proposes that the guidance on this technology is considered for review by the Guidance Executive 3 years after publication of the guidance. NICE welcomes comment on this proposed date. The Guidance Executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Amanda Adler
Chair, Appraisal Committee
June 2015

8 Appraisal Committee members, guideline representatives and NICE project team

Appraisal Committee members

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

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

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

Dr Amanda Adler (Chair)
Consultant Physician, Addenbrooke's Hospital

Professor Ken Stein (Vice Chair)
Professor of Public Health, University of Exeter Medical School

Dr Ray Armstrong
Consultant Rheumatologist, Southampton General Hospital

Dr Jeff Aronson
Reader in Clinical Pharmacology, University Department of Primary Health Care, University of Oxford

Professor John Cairns

Professor of Health Economics Public Health and Policy, London School of Hygiene and Tropical Medicine

Professor Imran Chaudhry

Lead Consultant Psychiatrist and Deputy Associate Medical Director, Lancashire Care NHS Foundation Trust

Dr Neil Iosson

Locum GP

Dr Sanjay Kinra

Reader in Clinical Epidemiology and Honorary Consultant in Paediatrics, London School of Hygiene and Tropical Medicine and University College London NHS Hospitals Trust

Dr Miriam McCarthy

Consultant, Public Health, Public Health Agency, Northern Ireland

Mr Christopher O'Regan

Head of Health Technology Assessment and Outcomes Research, Merck Sharp & Dohme

Dr Sanjeev Patel

Consultant Physician and Senior Lecturer in Rheumatology, St Helier University Hospital

Dr John Pounsford

Consultant Physician, Frenchay Hospital, Bristol

Dr Danielle Preedy

Lay Member

Mr Alun Roebuck

Consultant Nurse in Critical and Acute Care, United Lincolnshire NHS Trust

Dr Nicky Welton

Senior Lecturer in Biostatistics/Health Technology Assessment, University of Bristol

NICE project team

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

Martyn Burke

Technical Lead

Nicola Hay

Technical Adviser

Jeremy Powell

Project Manager

9 Sources of evidence considered by the Committee

A. The Evidence Review Group (ERG) report for this appraisal was prepared by CRD and CHE Technology Assessment Group, University of York:

- Simmonds M, Lomas J, Llewellyn A et al., Vortioxetine for treating major depressive disorder, April 2015

B. The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to make written submissions. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I. Company:

- Lundbeck

II. Professional/expert and patient/carer groups:

- Black Mental Health UK
- British Association for Psychotherapy
- College of Mental Health Pharmacy
- Depression Alliance
- Royal College of Psychiatrists

III. Other consultees:

- Department of Health
- NHS England
- Welsh Government

IV. Commentator organisations (did not provide written evidence and without the right of appeal):

- Cochrane Depression and Anxiety Group
- Department of Health, Social Services and Public Safety for Northern Ireland
- Health Improvement Scotland
- Merck Serono
- MRC Clinical Trials Unit
- Servier

C. The following individuals were selected from clinical expert and patient expert nominations from the consultees and commentators. They gave their expert personal view on vortioxetine by attending the initial Committee discussion and providing a written statement to the Committee. They are invited to comment on the ACD.

- Professor Heinz Grunze, Professor of Clinical Psychiatry, Academic Psychiatry and Regional Affective Disorders Service
- Newcastle University, nominated by Lundbeck – clinical expert
- Emer O'Neill, Chief Executive, Depression Alliance, nominated by Depression Alliance – patient expert

D. Representatives from the following company attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Lundbeck