NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Vortioxetine for treating major depressive episodes

Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)

Definitions:

Consultees – Organisations that accept an invitation to participate in the appraisal including the companies, national professional organisations, national patient organisations, the Department of Health and the Welsh Government and relevant NHS organisations in England. Consultees can make a submission and participate in the consultation on the appraisal consultation document (ACD; if produced). All non-company consultees can nominate clinical experts and/or patient experts to verbally present their personal views to the Appraisal Committee. Company consultees can also nominate clinical experts. Representatives from NHS England and clinical commissioning groups invited to participate in the appraisal may also attend the Appraisal Committee as NHS commissioning experts. All consultees have the opportunity to consider an appeal against the final recommendations, or report any factual errors, within the final appraisal determination (FAD).

Clinical and patient experts and NHS commissioning experts – The Chair of the Appraisal Committee and the NICE project team select clinical experts and patient experts from nominations by consultees and commentators. They attend the Appraisal Committee meeting as individuals to answer questions to help clarify issues about the submitted evidence and to provide their views and experiences of the technology and/or condition. Before they attend the meeting, all experts must either submit a written statement (using a template) or indicate they agree with the submission made by their nominating organisation.

Commentators – Commentators can participate in the consultation on the ACD (if produced), but NICE does not ask them to make any submission for the appraisal. Non-company commentator organisations can nominate clinical experts and patient experts to verbally present their personal views to the Appraisal Committee. Commentator organisations representing relevant comparator technology companies can also nominate clinical experts. These organisations receive the FAD and have opportunity to report any factual errors. These organisations include comparator technology companies, Healthcare Improvement Scotland any relevant National Collaborating Centre (a group commissioned by NICE to develop clinical guidelines), other related research groups where appropriate (for example, the Medical Research Council and National Cancer Research Institute); other groups such as the NHS Confederation, the NHS Commercial Medicines Unit, the Scottish Medicines Consortium, the Medicines and Healthcare Products Regulatory Agency, the Department of Health, Social Services and Public Safety for Northern Ireland).

Public – Members of the public have the opportunity to comment on the ACD when it is posted on the Institute's web site 5 days after it is sent to consultees and commentators. These comments are usually presented to the appraisal committee in full, but NICE reserves the right to summarise and edit comments received during consultations, or not to publish them at all, where in the reasonable opinion of NICE, the comments are voluminous, publication would be unlawful or publication would be otherwise inappropriate.

Please note: Comments received in the course of consultations carried out by NICE are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the submissions that NICE has received, and are not endorsed by NICE, its officers or advisory committees.

Comments received from consultees

Consultee	Comment [sic]	Response
Department of Health	No comments.	Comments noted. No action required.
Depression Alliance	Although there are a large number of antidepressants currently available in the UK, we need to see a wide range of efficacious and well tolerated treatments due to the heterogeneous nature of depression. Our members tell us that the potential of side effects and real experiences of side-effects hinder their willingness to try antidepressants but also adherence and persistence to continue with treatment. Additional well-tolerated products are needed to improve outcomes and quality of life associated with the treatment of depression. Based on our members' experiences, we want to see improved treatment options for issues such as cognition in depression which affect peoples' ability to both gain and retain employment.	Comments noted. Vortioxetine is recommended as an option for treating major depressive episodes in adults whose condition has responded inadequately to 2 antidepressants within the current episode (see FAD section 1.1).
Lundbeck	**Section 2: The Technology** *Section 2.2, Page 4*: 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. The adverse events listed above are no longer correct due to an update of the	Comments noted. The technology section has been updated in the FAD, see section 2.

SmPC. Please see Table 18 for the latest adverse events included in the SmPC for vortioxetine (EMA, 17th June 2015).

Table 18: Adverse reactions associated with vortioxetine

SYSTEM ORGAN CLASS	FREQUENCY	ADVERSE REACTION
Psychiatric disorders	Common	Abnormal dreams
Nervous system disorders	Common	Dizziness
	Unknown	Serotonin Syndrome
Vascular disorders	Uncommon	Flushing
Gastrointestinal disorders	Very common	Nausea
	Common	Diarrhoea, Constipation, Vomiting
Skin and subcutaneous tissue disorders	Common	Pruritus, including pruritus generalised
	Uncommon	Night sweats
Adverse reactions are listed using the following convention: very common (≥1/10); common (≥1/100 to <1/10); uncommon (≥1/1,000 to <1/100); rare (≥1/10,000 to <1/1,000); very rare (<1/10,000), not known (cannot be estimated from the available data).		

Section 2.3 Page 4: 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.

This is incorrect and should read as per the SmPC: The starting dosage is 10 mg

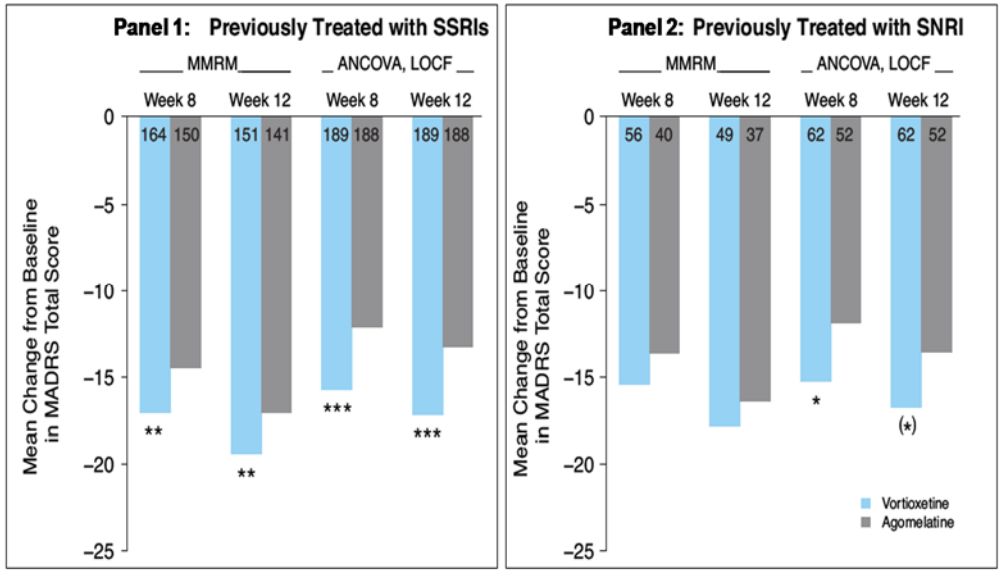
The technology section has been updated in the FAD, see section 2.

	once daily in adults younger than 65 years, and 5 mg once daily in adults 65 years and older. Depending on individual patient response, the dose may be increased to a maximum of 20 mg vortioxetine once daily or decreased to a minimum of 5 mg vortioxetine once daily (EMA, 2015. SmPC Section 4.2).	
	Section 3: The Company's submission *Section 3.3, Page 6 (& Section 4.6, Page 36): Full Analysis Set (FAS) vs Intention to Treat (ITT)* The ACD documents that Lundbeck presented results of the REVIVE trial using the Full Analysis Set (FAS) population rather than the Intention to Treat (ITT) population. A true ITT analysis is not straightforward since assumptions need to be made for patients without post-baseline data. For this reason FAS is generally accepted as a modified ITT. In the current study the problem is small with 3 and 5 ITT patients respectively in the VOR and AGO group not qualifying for FAS. Assuming e.g. a 'no change score' for these patients did not have impact on the results.	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, see section 4.6 of FAD.
	Section 3.3, Page 6: The company tested a primary hypothesis of non-inferiority, and a secondary hypothesis of superiority. This statement is incorrect. The superiority test was not a secondary hypothesis but part of the primary hypothesis. It is well recognised that once non-inferiority is established in a non-inferiority study a superiority test can be performed under a closed test procedure (EMA 2000). Non-inferiority was established in REVIVE as the upper bound of the 95% CI for the vortioxetine and agomelatine comparison was -0.81 MADRS points. This was clearly below the non-inferiority margin of +2 MADRS points vs. agomelatine. As the two-sided 95% CI for the difference between the means excluded zero and was in favour of vortioxetine, vortioxetine was superior to agomelatine (mean difference for vortioxetine to agomelatine was 2.2 ({95% CI: 3.5 to 0.8; p = 0.0018}).	Please note section 3.3 of the FAD has been updated.
	Section 3.5, Page 7: 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).	Please note that section 3.5 of the FAD has been updated.

	This sentence is factually inaccurate. It should read as follows: The company stated that vortioxetine statistically significantly improved outcomes compared with agomelatine across the analyses of response and remission outcomes measured by MADRS score (see table 2).	
	Section 3.9 Page 9: “There are no head-to-head data comparing vortioxetine with comparators other than agomelatine in the second-line population.”	Please note that section 3.9 of the FAD has been updated.
	This statement is incorrect. As stated in the submission there were two trials of vortioxetine in switch populations though these studies were dissimilar in design and methodology. Study 14178A (REVIVE) was an efficacy study in patients switching due to lack of efficacy, while Study 318 was a tolerability study in patients who were well-treated but required a treatment switch due to experiencing TESD with their current treatment. Study 14178A (REVIVE) compared vortioxetine with agomelatine and Study 318 compared vortioxetine with escitalopram.	
	In addition to REVIVE and TAK318, a head-to-head study vs venlafaxine has also been conducted in patients with MDD. Further information about the study is provided in Appendix 4.46.3 and the results of this study are included in the cost-effectiveness analyses undertaken and reported in Section 3.7).	
	Section 3.14, Page 14 (& Section 3.36, Page 24; Section 4.16, Page 47) 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).	No action required.
	Though Lundbeck does appreciate that a half cycle correction is usually applied to both the costs and health effects the rationale for not applying the half-cycle correction for the former was as follows: - Resource use and costs were available for the period preceding the health state achieved after the acute phase, e.g. for remission the cost is available for the two-month period before achieving remission. This principle was applied for all cycles, i.e. in the Markov phase, the cost of the acute phase corresponds to the two-month	

	cost associated with achieving remission or of remaining in non-remission depending on the outcome. - Quality of life/utilities were available at the time of the health states, e.g. at 8 weeks utilities were available when patients should a patient achieve remission or not. Therefore to account for the transition between health states at different points in time, half-cycle correction was applied. As stated above, despite the rationale employed leading to the half-cycle correction to be applied to utilities and not costs it is recognised that a half-cycle correction is applied to both costs and outcomes. Therefore all additional analyses provided in this document apply a half-cycle correction to both costs and utilities. *Section 3.23, Page 19 (& Section 4.5, Page 35):* The ERG commented that the population enrolled into REVIVE was broadly representative of the second-line population in England… 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… Lundbeck do not believe that either of these factors lessen the external validity of REVIVE. As stated in the submission, the efficacy of vortioxetine did not differ by race in subgroup analyses of the full MDD package and therefore should not affect the generalisability of results (vortioxetine EPAR, 2013). In addition, one of the subgroup analyses presented within the Manufacturer's submission was by previous class of treatment (SSRI or SNRI). For both vortioxetine and agomelatine, the mean MADRS total score decreased (improved) throughout the 12-week treatment period both for patients previously treated with either class of antidepressant (Papakostas 2014).	Section 3.23 is from the view of the independent Evidence Review Group and is not considered a factual inaccuracy. No action required.

	For patients previously treated with an SSRI, the differences between vortioxetine (n=164) and agomelatine (n=150) at Week 8 on MADRS total score (FAS, MMRM) was -2.6 (95% CI: -4.1 to -1.0; p=0.0013). The robustness of the results was shown by sensitivity analyses using ANCOVA (FAS, LOCF), where the mean difference to agomelatine at Week 8 on MADRS total score was -3.6 (95% CI: -5.4 to -1.9; p<0.0001). Similar results were seen at Week 12 (Figure 21, Panel 1). For patients previously treated with an SNRI, the differences between vortioxetine (n=56) and agomelatine (n=40) at Week 8 on MADRS total score (FAS, MMRM) was -1.8 (95% CI: -4.8 to 1.2; p=0.2254) (Figure 2). For the sensitivity analyses using ANCOVA (FAS, LOCF) the mean difference to agomelatine at Week 8 on MADRS total score was -3.4 (95% CI: -6.6 to -0.1; p=0.0420). Similar results were seen at Week 12 (Figure 21, Panel 2). It can be concluded that the efficacy advantage of vortioxetine over agomelatine is independent of previous treatments class (SSRI or SNRI). Therefore the SNRI pre-treatment of 23% of the population in REVIVE does not affect the generalisability of results. Figure 21: Study 14178A (REVIVE). Subgroup analysis of primary efficacy endpoint by previous class of treatment	



Source: Poster (Papakostas 2014). Derived from subgroup analysis reported in CSR27, various tables. FAS: full-analysis set, LOCF: last observation carried forward, MMRM: mixed model, repeated measures. Patient numbers at each visit are shown below the x-axis for each treatment group. (*) p=0.06; *p<0.05; **p<0.01; ***p<0.001 versus agomelatine.

Section 3.25, Page 19: 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.

Lundbeck do not agree that the confidence intervals around the results observed in REVIVE should be considered "relatively wide". A MADRS SD=7.6 was observed in the study which is lower than what is typically found in the other vortioxetine studies in MDD (SD=9.5). In addition with n~220 per group these confidence intervals are in fact narrower than what is typically found.

Section 3.27, Page 20 (& Section 4.7, Page 37): It also stated that time of outcome

Section 3.25 is from the view of the independent Evidence Review Group and is not considered a factual inaccuracy. No action required.

Section 3.27 is from the view of the independent Evidence Review Group and is not considered a

	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. As stated in the submission Lundbeck does not believe the variation in time points of outcome assessment between trials affects the results of the indirect comparison in switch. Allowing a range is common practice in data syntheses and was not expected to be an effect modifier with respect to relative effectiveness. Justification is provided by the findings of REVIVE (REVIVE CSR 2013), in which the difference between vortioxetine and agomelatine in change from baseline of the total MADRS score was almost constant at the week 4, week 8 and week 12 (end of study) assessment points (see Manufacturer's submission, Section 6.10.2.3). The time-point for reporting withdrawals due to AEs also differed between studies. However, this again was not expected to affect relative treatment effects greatly as most withdrawals due to AEs occur within 2-4 weeks of starting treatment (NHS Choices, 2013 & REVIVE analyses presented in Section 3.6.2.4).	factual inaccuracy. No action required.
	Section 3.33, Page 22 (& Section 4.12, Page 41): 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 current model design was selected based on both expert recommendations and previously published cost-effectiveness models in MDD. Both sources used to inform the structure highlighted the high heterogeneity between patients and patient pathways in both the treatment and management of the disorder. A decision tree was selected to reflect the different clinical (i.e. outcome) pathways a patient can have following the acute and maintenance phases of treatment. In order to reflect clinical practice patients could switch at different points in time depending on the reason (e.g. treatment discontinuation due to relapse or long-term adverse events). Therefore Lundbeck believed it was appropriate a to start the Markov	Section 3.33 is from the view of the independent Evidence Review Group and is not considered a factual inaccuracy. No action required.

	component at the time of switch (rather than, for example, after the first 8 weeks of treatment) and adjust the time spent in the Markov up to the end of the time horizon. **Section 4: Consideration of the evidence** *Section 4.3, Page 33*: 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. Lundbeck support the clinical expert's view that the treatment of depression should consider aspects such as treatment history and an individual's symptoms as depression is a heterogeneous disorder. This is in line with the recommendations in CG90 which states that antidepressants are largely equal in efficacy therefore choice should be primarily dependant on safety, tolerability and patient preference. However, prescribing data presented within the Manufacturer's submission (again presented in this document Figure 1) shows escitalopram not to be the most commonly used antidepressant at second line within the NHS. *Section 4.6, Page 36*: 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. Lundbeck believes the clinical experts view has been misreported within the ACD. Firstly, although Lundbeck recognise that the MADRS scale is not widely used in clinical practice, it is however recognised as a standard measurement of symptoms associated with depression, in clinical research, alongside HAM-D (Duru & Fantino 2008).	No action required. The Committee was aware that the company's evidence included people with first and recurrent major depressive episodes. It recognised that it was appropriate to change treatments between the first episode and any subsequent episodes, and to not use any drug, or specific sequence of drugs, for a recurrent episode that has not worked during a previous episode. The Committee also noted that the NICE's guideline on depression in adults states that 'treatment choice should be influenced by: previous treatment history, including the consequences of a relapse, residual symptoms, response to previous treatment, any discontinuation symptoms, and the person's preference'. The Committee therefore agreed that a flexible approach was needed in clinical practice when treating recurrent episodes of depression.' Please see section 4.20 of the FAD. The Committee noted that this was disputed in some comments it received on the appraisal consultation document. However, 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. For further information, please see section 4.6 of the FAD.

	The Committee heard the clinical expert state that, although there is some debate as to what constitutes a clinically meaningful change in MADRS score it is generally accepted that 2 points are considered the threshold in a placebo-controlled trial. Within active-controlled trials this difference can be halved and is still considered clinical meaningful. This is in line with the literature provided by Lundbeck at the clarification question stage of the appraisal process which included the following. Table 19 details what is considered a clinically meaningful change in symptom scores for the relevant scales as detailed in the literature. These clinically relevant differences are however in the context of assessing pharmacological compounds versus placebo. As discussed by Montgomery (2011), this clinically relevant difference versus placebo is not to be generalised for comparisons between two active drugs, and one rule that can be applied for the translation from placebo-control to active-control comparison is 50% of the difference judged as clinically relevant in placebo-controlled studies: “The criteria used to establish a clinically relevant difference have almost all been focused on a comparison of drug and placebo. Comparing differences between two active treatments applying the same criterion used to define a clinically relevant difference on the pivotal scale between active drug and placebo is very stringent, since this means that the difference to placebo of the superior treatment must be at least twice that of the comparator antidepressant. However, 50% of the difference between active drug and placebo has also been used as a criterion to indicate probable clinical relevance when comparing two established treatments.” The London New Drugs Group (LNDG) has accepted this criterion when assessing evidence to inform local decision-making (Denby 2009). This is demonstrated in their discussion of agomelatine and what should be considered a clinically relevant change in HAM-D17 score: “The agomelatine/placebo difference in HAMD score was not always greater than 3 points (ref 8 and the active control studies), the difference considered to be clinically	

significant by NICE. The agomelatine/sertraline HAMD score difference was greater than 1.5 points, which is considered clinically significant."

Table 19: Clinical relevance of changes in symptom scores on depression rating scales

Scale	Clinically relevant difference in change in symptom score: active treatment versus placebo control		References
		Half of the score considered clinically relevant in active-controlled vs. placebo controlled studies based on Montgomery 2011[4]	
MADRS	2 points	1 point	Batterham 2009 Melander 2008 Bruce 1994
	1.6 to 1.9 points	0.8 to 0.95 points	Duru 2008
HAM-D_{24}	2 points	1 point	Duru 2008 Montgomery 2009
HAM-D_{17}	3 point	1.5 points	Denby
CGI-S/CGI-I	No clear consensus		Bandelow 2006

The clinical expert did continue to state that at an individual patient level he would be able to observe a reduction in 1 item of the MADRS by 2 or more points.

Section 4.8, Page 38: 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

Section 4.8 of the FAD has been updated.

	relativeness effectiveness changes across lines of treatment. There is a typographical error in this sentence. “Relativeness” should be replaced with “relative”. *Section 4.8, Page 38*: 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. Lundbeck conducted a systematic review to identify all data available to answer the research question of whether there are differences in relative treatment effects at different treatment lines. The full report can be found in Appendix 6.2. The conclusions from this review were as follows: “The systematic review demonstrates that there is limited evidence currently available allowing the assessment of relative efficacies of antidepressants across treatment lines.” Despite identifying a number of studies providing some evidence that may be used to interpret the difference in relative efficacy of antidepressants between lines of treatment, quantitative analysis would be extremely difficult to perform with the reported data. The Bauer publication (16) potentially provides the strongest evidence in answering the question of whether relative efficacies differ between treatment lines. Despite limitations associated with the design, as each treatment arm used as evidence consisted of at least 2,000 patients from several clinical trials, these can be considered robust results. Therefore, the reported odds ratio patterns may be indicative of a variation in relative treatment efficacy between the before-and after-	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 the company had not provided sufficient evidence, either in its original submission or in its response to the appraisal consultation document, to support a declining relative effect of treatment between non-SSRIs within each line of treatment. For further information, please see section 4.8 of the FAD.

	switch populations. Considering the evidence presented in Figure 13 in its totality, it appears that the general trend towards a difference in the number of patients achieving a response to treatment between before- and after-switch populations between interventions. The findings of this review add to the understanding of relative efficacies gained from the findings of the Papakostas (1) publication; that across-class switch from SSRI to non-SSRI is more likely to result in remission than switching from one SSRI to another". *Section 4.9, Page 39*: 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. Lundbeck does not believe Llorca (2014) to provide the most robust evidence for populating the current cost-effectiveness model. See Section 3.4 for full rationale. *Section 4.10, Page 40*: 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 effect of vortioxetine in the switch population has been demonstrated in a trial longer than 8 weeks - the REVIVE study which was 12 weeks in duration. *Section 4.11, Page 40*: The Committee understood that the dose of vortioxetine generally increases over time. Lundbeck are unclear on the evidence base used to substantiate this. Within the REVIVE study the starting dose for all patients randomised to vortioxetine was 10mg. This could be adjusted (either to 10mg or 20mg) at the weekly visits during	The Committee was aware from the response to the ACD that the company was concerned that the results from Llorca for remission and response were potentially biased because several of the included trials did not present data for these outcomes. The Committee heard from the ERG that there was no evidence to suggest that the Llorca analysis was affected by reporting bias. 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. For further information, please see sections 3.48 and 4.9 of the FAD. Please note that sections 4.10 and 4.11 of the FAD have been updated.

	the first 4 weeks only. In clinical studies with flexible-dose design there is a tendency to increase the dose. In REVIVE, the proportion of patients on the high dose (approximately 2/3 of the patients) is in line with what has been observed in the clinical development programme for vortioxetine for other studies with a flexible-dose design (Study 11985A, Table 36 [open-label period], and open-label extension studies [Study 11492C, chapter 8, Study 11984B, chapter 8]). In REVIVE the investigator could only adjust the dose within the first 4 weeks of the treatment period, which might have affected the titration and the speed of up-titration as compared to what could be expected in clinical practice. This phase IIIb study included a common and difficult-to-treat population than the studies in the MAA did. Being aware of this, and of the dose-response relationship of vortioxetine, the investigators may have tended to use higher doses in this particular population. In addition, it is difficult to fully translate the observed dose pattern in these studies to what could be expected in clinical practice. Based on the clinical trial data currently available it could be said that rather than the dose “generally increases over time” the likelihood is that a more substantial proportion of patients are up-titrated to the high dose. Again, whether this reflects reality is another question as it is often seen that within clinical trials, doses are more readily up-titrated than in usual clinical practice. *Section 4.13, Page 44*: 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). Lundbeck does not support the view that patients failing first-line antidepressant	Please see FAD sections 4.3 (general context) and 4.4 (specific to vortioxetine) which state: “The Committee was aware that the NICE’s guideline depression in adults gave GPs the option to prescribe second-line treatments in primary care (for example, escitalopram or an SNRI) and that GPs may also manage people for whom third-line treatment is needed. The Committee understood that clinicians would like to use vortioxetine for people whose major depressive episode is likely to benefit from second- or third-line treatment (that is,

	treatment are referred to secondary care or a community mental health trust is representative of clinical practice in the majority of the UK. In line with NICE guidelines, the insight gained from clinicians both in primary, secondary and tertiary care is that in the main patients with depression are referred after inadequate response to 2-3 antidepressants in primary care. The exception to this is if a patient is suicidal, complex or has depression with psychotic symptoms. In these cases the patient would be automatically referred for specialist care (NICE 2009 and advisory board Appendix 6.4).	after SSRI therapy) with a 'newer-generation, better tolerated antidepressant'."
	Section 4.21, Page 51: The Committee concluded that these benefits were sufficiently captured within the company's economic modelling. Lundbeck believes the Committee to have overlooked the benefits associated with the lack of discontinuation symptoms experienced with treatment with vortioxetine. All antidepressants are associated with discontinuation symptoms bar agomelatine and vortioxetine (Taylor et al. 2015). As stated in the submission poor adherence to antidepressant medication is common and patients are often seen to miss consecutive doses of treatment for several days (Demyttenaere et al. 2001). Haddad and Anderson (2007) note that such breaks in therapy can precipitate discontinuation symptoms. In their review of all antidepressants, they reported that the most common discontinuation symptoms included dizziness, headache, nausea and lethargy. Mostly these symptoms were short-lived and mild but in some cases they can be severe, lasting several weeks. As discontinuation symptoms are not included within the QALY calculation due to disutilities of such side-effects being unavailable from the literature, the cost-effectiveness of vortioxetine may be underestimated versus treatments associated with such symptoms.	No action required.
	Section 4.21, Page 52: The Committee acknowledged that vortioxetine may be a valuable treatment option for people with a major depressive order experiencing cognitive dysfunction.	Section 4.21 of the FAD has been updated.

There is a typographical error in this sentence. The sentence should read:

"The Committee acknowledged that vortioxetine may be a valuable treatment option for people with a major depressive disorder experiencing cognitive dysfunction"

Section 4.21, Page 52: The Committee…. 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.

This statement is not correct. Statistically significant differences in EQ-5D scores in favour of vortioxetine over agomelatine were observed in REVIVE. As reported in the submission the EQ-5D profile data were scored on a summary index, using the UK tariff (Dolan 1997); summary statistics of index scores at each assessment point are shown in Table 20. The mean EQ-5D summary index increased in both treatment groups over the 12-week treatment period from 0.53 (vortioxetine) and 0.55 (agomelatine) at baseline to 0.81 (vortioxetine) and 0.78 (agomelatine) at Week 12 (FAS, OC).

Table 20: Study 14178A (REVIVE). EQ-5D Summary index scores (FAS, OC)

Assessment	Vortioxetine			Agomelatine		
	N	Mean	sd	N	Mean	sd
Baseline	252	0.53	0.28	241	0.55	0.27
Week 4	241	0.70	0.22	233	0.64	0.27
Week 8	220	0.76	0.19	189	0.73	0.23
Week 12	200	0.81	0.21	178	0.78	0.22

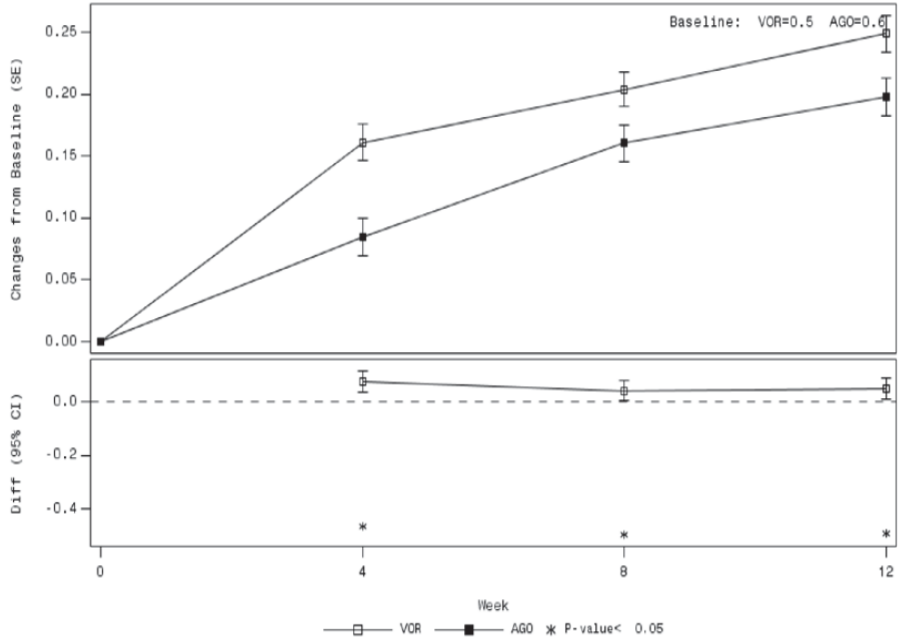
Source: REVIVE CSR Table 195. Sd: standard deviation; FAS: full analysis set. OC: observed cases. sd: standard deviation

When comparing changes in utility score from baseline between the two groups, greater improvements were seen in the vortioxetine group than in the agomelatine

Section 4.21 of the FAD has been updated. The Committee noted that the EQ-5D data from REVIVE reported for the vortioxetine and agomelatine groups did not suggest that the average utility for remission and non-remission was notably different between treatments.

group from week 4 onwards, as shown in Figure 22. At week 4, the changes from baseline scores were 0.16 and 0.08 for the two groups, respectively, (p<0.001); at week 8, 0.20 and 0.16 (p =0.03) and at week 12, 0.25 and 0.20 (p =0.01) (FAS, MMRM). Slightly greater differences were seen in the sensitivity analysis using ANCOVA (FAS, LOCF).

Figure 22: Study 14178A (REVIVE). Mean changes from baseline in EQ-5D summary index and mean treatment differences to agomelatine (FAS, MMRM)



Source: CSR, Panel 33. FAS: full analysis set; MMRM: mixed model for repeated measures.

Quality-adjusted survival over 12 weeks was calculated for each study group using the areas under the utility curves from each assessment point (see Figure 22). The area between these curves thus represents the mean quality-adjusted survival

	difference between the two treatments over 12 weeks of treatment.	
SANE	**The importance of choice for patients** There is a paucity of choice for people suffering from severe depression. This is a condition which can have debilitating negative effects, for example on cognition, which can affect not only an individual's quality of life but the ability to work, and can cause some people to be at risk of suicide. We are told by those contacting us that despite all the promises made about choice of treatments, patients are still given very little choice in medication. Greater choice and more tolerable side effects for the individual than those often associated with existing medications are of paramount importance to a person undergoing a major depressive episode. The side effects of existing drugs can limit a patient's willingness to take medication, compromise adherence, and inhibit persistence in continuing with the treatment. SANE is not in a position to judge the clinical efficacy and cost-effectiveness of any one medication. We would argue strongly, however, that any new drug using different chemical pathways which offered relief from the negative effects of major depressive episodes, in particular on cognition, and had better-tolerated side-effects, would be important in improving treatment outcomes and be more than welcome to patients. We believe that any medication which adds to the armoury of treatment available is worth a close study of the benefits to patients and the risks associated with its not being available as a potential treatment. If Vortioxetine were proven to be a clinically effective treatment for major depressive episodes in adults, and cost-effective, it would enable doctors to offer a wider choice of medications and could have a transformative effect on some patients.	Comments noted. The Committee concluded that vortioxetine could be recommended as an option for treating major depressive disorder in adults whose condition has responded inadequately to 2 antidepressants within the current episode (see FAD section 1.1).

Comments received from commentators

Commentator	Comment [sic]	Response
Cochrane Depression, Anxiety and Neurosis Review Group (CCDAN)	We would like to make some comments about the clinical efficacy part of the document. We have not commented on the cost-effectiveness analysis. **General comments** The company submitted evidence for the second line population only, so there is little overlap with our Cochrane review. NICE have taken account of previous reviews, one of which has a set of studies similar to our Cochrane review - we identified an additional study comparing vortioxetine with venlafaxine, and two studies with duloxetine. This evidence was not submitted by the company. The other review considered by NICE is an indirect comparison meta-analysis (http://www.ncbi.nlm.nih.gov/pubmed/25249164) sponsored by the company. **1. The appropriateness of agomelatine as comparator in a non-inferiority trial** In light of our Cochrane review, we think the choice of agomelatine in a non-inferiority trial in the submission is questionable, and we have some concerns about the company submission with respect their justification for this. Their submission reads: “Agomelatine: although not widely used within the UK it is recognised as an efficacious and well-tolerated antidepressant that has a mode of action distinct from the SSRIs and SNRIs87. The most recent NICE surveillance review for CG90101 identified two meta-analyses31,102, “which indicate that agomelatine is at least as efficacious as escitalopram, fluoxetine, sertraline, venlafaxine and paroxetine in reducing depression scores for response and remission in patients with depression and severe depression”. (page 47) and “Although agomelatine is seldom used within the NHS, a recent independent meta-analysis found it to be effective164. Also, it is commonly prescribed in other	Comments noted. 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. Please see section 4.5 of the FAD.

	European countries and fulfils the criteria for a “better tolerated, newer generation” antidepressant within the NICE treatment pathway, according to expert clinical advisors consulted by Lundbeck” (page 65) Our Cochrane review shows, that paroxetine may be better than agomelatine. The effect was not significant, but the SMD effect size is similar to the one from our comparison of agomelatine vs. placebo. Furthermore, in the light of the liver problems with agomelatine, the assertion that agomelatine is “better tolerated” seems not to be justified. In our agomelatine review, we conclude that the efficacy and clinical utility of agomelatine is rather low. The poor efficacy and tolerability data for agomelatine makes it a very surprising choice of drug to provide a fair test in a non-inferiority trial in the submission, and we would question the company submission in relation to the appropriateness of this drug. **2. Concerns about the adequacy of the network meta-analysis.** We recognize that the ERG partly acknowledge the weaknesses in the network, but we feel that the extent of the limitations and the impact of these is insufficiently clear. Vortioxetine is only linked to the network via one study Kasper (2013). The interpretation of the network data should be much more cautious. The consultation document states: “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.” (page 20, point 3.26) The York ERG report reads: “It was unclear whether all such patients were genuinely switching patients, or	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 in Kasper may not be comparable with the other populations in the evidence network, or consistent with the population specified in its decision problem. 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 second-line treatment. For further details, please see section 4.7 of the FAD.

	whether they had been treated for a prior depressive episode. The ERG considers that the patients are likely to be a mix of these cases, and therefore the eligibility of these two trials for an indirect treatment comparison of switch populations is questionable." (page 60.) The company submission reads in their limitations: "For instance, Kasper (2013) provides the data for agomelatine vs. sertraline and placebo (Kasper 2010, Olié and Kasper 2007, respectively). In this publication, a post-hoc analysis was undertaken which provided the data for the "previously treated" population. This population was defined as patients receiving any antidepressant treatment within the previous 12-month period. This criterion may have allowed patients who were not necessarily switching antidepressants within the current episode to be included in the study. Therefore the population in Kasper 2013 may be slightly wider than those of the other trials. A conclusion on the impact on the estimates of relative effects cannot be drawn. In such a post-hoc analysis, randomisation is not strictly preserved, but expert health economists and statisticians considered this to be a theoretical rather than practical consideration and therefore did not expect it to be an effect modifier." (page 137) These statements require clarification, and a more detailed examination of the studies referred to yields greater clarity and a potentially different interpretation. The subgroup included from the two agomelatine trials (Kasper 2010 and Olie & Kasper 2007) had some antidepressant treatment in the last 12 month, but do not refer to non-response anywhere in their inclusion criteria. In fact, Kasper 2010 (which is the main comparison) excluded patients with "previous depression resistance to antidepressants" (Kasper 2010, page 111) and Olie & Kasper excluded "previous resistance to antidepressants (defined as the failure to respond to two different prior antidepressant medications prescribed for >=4 wk during the current depressive episode)" (page 663). In addition, both trials used wash-out periods, which lasted up to 5 weeks if the patients were treated with fluoxetine or trazodone in the Kasper 2010 study.	

	We would NOT regard this as a switching population. Switching would have been the exception rather than the rule in this group of patients. According to Star*D there is a decline in the response rates in every treatment step (as noted in the company submission), so the agomelatine response rates seem to be clearly overestimated for second line treatment. This is of major importance, because vortioxetine is only linked to the network via the Kasper study (see page 116 in the company submission). We would therefore question the validity of the entire network. Lastly, the evidence for agomelatine as a second line treatment in treatment resistant depression is weak. To our knowledge, the only trial of agomelatine in treatment resistant depression was a pilot study by Servier (CL3-027), which remained unpublished. **3. Clinically important outcome** We do not agree with their clinical expert, that "The clinical expert explained that a reduction in 1 item of the MADRS by 2 or more points would generally be considered clinically meaningful". (page 36) We recognize this is of minor importance, as the Committee agreed that Remission and relapse preventions are more useful to judge clinical efficacy. **4. Error** On page 20 (see 3.26) the document states that the Rosso trial was the only trial that compared vortioxetine with duloxetine, but the trial compared duloxetine and bupropion.	The Committee noted that this was disputed in some comments it received on the appraisal consultation document. However, 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. For further details, please see section 4.6 of the FAD. Section 3.26 of the FAD has been updated.
Merck Serono	No comments.	Comments noted. No action required.
Servier Laboratories Limited	Overall Servier's view is that general the document is balanced and fair in its assessment. Point 3.29 "The ERG considered that Llorca may represent the most reliable	Comments noted. No action required.

	evidence for comparing vortioxetine with other treatments." Servier acknowledge this may be true, however wish to highlight that Taylor 2014 is another non-industry sponsored meta-analysis which provides a relevant assessment of the efficacy of agomelatine compared to other comparators in the current appraisal (ID583). Reference: Taylor D, Sparshatt A, Varma s, Olofinjana 0. Antidepressant efficacy of agomelatine: meta-analysis of published and unpublished studies. BMJ 2014; 348:g1888.	

Comments received from members of the public

Role*	Section	Comment [sic]	Response
Healthcare professional (within NHS)	General	The majority of people with major Depressive disorder (MDD) are seen and managed in Primary Care. When considering treatment failures it is important to differentiate lack of efficacy from tolerability issues. It is only possible to assess potential for efficacy if a patient is able to tolerate therapy for a minimum of two weeks; response will usually be seen between 4 and 8 weeks and remission 8 to 12 weeks or longer. One or two changes of antidepressant may be required before a drug the patient can tolerate is found. Then further changes in dose or drug may be required for optimal efficacy. People referred to secondary care usually fall into the following categories - evidence of psychotic symptoms, other AXIS 1 co-morbidity, evidence of suicidality, failed efficacy with / without tolerability issues. For those patients with ongoing tolerability issues without other concerns Primary Care often ask advice rather than refer. Of those reviewed in secondary care many are treated and referred back to Primary care with advice about ongoing management. Only those on combinations of antidepressants, or antipsychotics, those at risk of suicide	Comments noted. Please see sections 4.3 (general context) and 4.4 (specific to vortioxetine) of the FAD which state: "The Committee was aware that the NICE's guideline depression in adults gave GPs the option to prescribe second-line treatments in primary care (for example, escitalopram or an SNRI) and that GPs may also manage people for whom third-line treatment is needed. The Committee understood that clinicians would like to use vortioxetine for people whose major depressive episode is likely to benefit from second- or third-line treatment (that is, after SSRI therapy) with a 'newer-generation, better tolerated antidepressant'."

* When comments are submitted via the Institute's web site, individuals are asked to identify their role by choosing from a list as follows: 'patent', 'carer', 'general public', 'health professional (within NHS)', 'health professional (private sector)', 'healthcare industry (pharmaceutical)', 'healthcare industry'(other)', 'local government professional' or, if none of these categories apply, 'other' with a separate box to enter a description.

Role*	Section	Comment [sic]	Response
		are likely to be retained in secondary care. For the most part patients seen in secondary care are quite different from those seen in Primary Care.	The Committee heard from the clinical expert that people with major depressive disorder whose condition responds after 8 to 10 weeks of treatment, but does not remit, would generally be treated for a further 4 weeks. The Committee also noted a comment received on the appraisal consultation document, which highlighted that people treated for major depressive disorder in primary care were likely to be completely different from those treated in secondary care. For further details around the Committee's deliberations of these aspects in relation to the company's economic modelling, please see section 4.13 of the FAD.