CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report
Vortioxetine for treating major depressive disorder

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD

Authors Mark Simmonds, Research Fellow, CRD

James Lomas, Research Fellow, CHE

Alexis Llewellyn, Research Fellow, CRD

Marta Soares, Senior Research Fellow, CHE

Kath Wright, Information Specialist, CRD

Alison Eastwood, Senior Research Fellow, CRD

Stephen Palmer, Professor, CHE

Correspondence to Prof. Stephen Palmer, Centre for Health Economics, University of York, York YO10 5DD

Date completed 13/04/2015

Source of funding

This report was commissioned by the NIHR HTA Programme as project 12/66/001

Declared competing interests of the authors

None

Acknowledgements

We thank Paul Blenkiron, Consultant Psychiatrist, Bootham Park Hospital, York, and Catherine Snape, General Practitioner, Jorvik Gillygate Practice, York, who acted as clinical expert advisors to the ERG.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Simmonds M, Lomas J, Llewellyn A, Soares M, Wright K, Eastwood A, Palmer S. Vortioxetine for treating major depressive disorder: A Single Technology Appraisal. CRD and CHE Technology Assessment Group, 2015.

Contributions of authors

Mark Simmonds and Alexis Llewellyn wrote the clinical effectiveness sections of the report. James Lomas and Marta Soares wrote the cost effectiveness sections of the report and performed the economic modelling. Kath Wright wrote the sections of the report dealing with search strategies and provided information support. Alison Eastwood was an advisor for the clinical effectiveness sections and commented on drafts of the report. Stephen Palmer wrote the critique of manufacturer's definition of decision problem, was an advisor for the cost effectiveness sections, and commented on drafts of the report.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in blue and underlined, all academic-in-confidence (AIC) data are highlighted in yellow and underlined

Table of Contents

List of abbreviations 9
1 Summary 10
1.1 Critique of the decision problem in the manufacturer's submission 10
1.2 Summary of clinical effectiveness evidence submitted by the manufacturer 11
1.2.1 Direct evidence in the switch population 11
1.2.2 Indirect evidence in the switch population 13
1.2.3 Evidence syntheses of non-switch populations 13
1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted 15
1.4 Summary of cost effectiveness submitted evidence by the manufacturer 17
1.4.1 Evidence used to inform the decision model 18
1.5 Summary of the ERG's critique of cost effectiveness evidence submitted 19
1.6 ERG commentary on the robustness of evidence submitted by the manufacturer 20
1.7 Weaknesses and areas of uncertainty 20
1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG 21
1.9 Conclusions from ERG analyses 22
2 Background 23
2.1 Critique of manufacturer's description of underlying health problem. 23
2.2 Critique of manufacturer's overview of current service provision 24
3 Critique of manufacturer's definition of decision problem 27
3.1 Population 27
3.2 Intervention 30
3.3 Comparators 32
3.4 Outcomes 37
3.5 Other relevant factors 38
4 Clinical Effectiveness 39
4.1 Critique of the methods of review(s) 40
4.1.1 Searches 40
4.1.2 Inclusion criteria 41
4.1.3 Critique of data extraction 42
4.1.4 Quality assessment 42
4.1.5 Evidence synthesis 42
4.2 Critique of the trials of vortioxetine 44
4.2.1 Critique of the validity of the included trials 44
4.2.2 Quality assessment and risk of bias 45
4.2.3 Characteristics of the REVIVE trial 47

4.2.4 REVIVE trial results 49
4.2.5 Characteristics of the TAK318 trial 53
4.2.6 TAK318 trial results 55
4.2.7 Conclusions from the review of vortioxetine RCTs 57
4.3 Indirect treatment comparison 58
4.3.1 Critique of trials identified and included in the indirect comparison 58
4.3.2 Critique of the network meta-analysis 61
4.3.3 Conclusions of the indirect treatment comparison 65
4.4 Non-RCT evidence 65
4.5 Adverse events 65
4.5.1 Short-term pool of placebo controlled studies 66
4.5.2 Open-label pool of continuation studies 69
4.5.2.1 Conclusions of the review of safety studies 71
4.6 Evidence syntheses of non-switch populations 72
4.6.1 Vortioxetine vs placebo 72
4.6.2 Vortioxetine vs active comparators 74
4.6.2.1 Pae et al. 74
4.6.2.2 Meeker et al. 76
4.6.2.3 PBAC submission 76
4.6.3 Indirect treatment comparison 77
4.7 Conclusions of the clinical effectiveness section 78
5 Cost Effectiveness 82
5.1 ERG comment on manufacturer's review of cost-effectiveness evidence 83
5.1.1 Searches 83
5.1.2 Inclusion/exclusion criteria used for study selection 83
5.1.3 Studies included and excluded in the cost effectiveness review 84
5.1.4 Conclusions of the cost effectiveness review 85
5.2 The manufacturer's economic evaluation compared with the NICE reference case checklist 89
5.3 ERG's summary and critique of manufacturer's submitted economic evaluation 90
5.3.1 Population 90
5.3.2 Model structure 90
5.3.2.1 Response and remission 95
5.3.3 Interventions and comparators 100
5.3.4 Perspective, time horizon and discounting 100
5.3.5 Treatment effectiveness and extrapolation 101
5.3.6 Relapse 104

5.3.7 Adverse events and withdrawals due to adverse events 104
5.3.8 Health related quality of life 106
5.3.9 Resources and costs 109
5.3.9.1 Overview 109
5.3.9.2 Acquisition costs 109
5.3.9.3 Health state costs 111
5.3.10 Discounting 113
5.3.11 Model validation and face validity check 113
5.3.12 Cost effectiveness results 115
5.3.12.1 Results from MS base case 115
5.3.12.2 Results from MS sensitivity analysis 115
5.3.12.3 Results from MS clarification scenarios 120
5.4 Conclusions of the cost effectiveness section 122
6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG 126
6.1 Overview 126
6.2 ERG initial corrections and adjustments to the manufacturer's base case model 127
6.3 Additional ERG analyses 129
6.3.1 Considering the broader effectiveness evidence 129
6.3.2 Examining alternative assumptions regarding the outcomes with further lines of therapy 135
6.4 Conclusions from ERG analyses 138
7 End of life 140
8 Overall conclusions 141
8.1 Clinical effectiveness 141
8.2 Cost effectiveness 142
8.3 Strengths, weaknesses and areas of uncertainty 144
8.3.1 Strengths 144
8.3.2 Weaknesses 144
8.3.3 Areas of uncertainty 145
8.4 Implications for research 146
9 References 147

Table of Tables

Table 1 RCTs included in the submission 11
Table 2 Summary of the results of the frequentist network meta-analyses 13
Table 3 Indirect treatment comparison by Llorca et al 14
Table 4 MS base-case results 19
Table 5 Alternative scenarios regarding relative effectiveness: cost-effectiveness (with up-titration) 21
Table 6 Summary of vortioxetine 31
Table 7 Comparators outlined in NICE scope 33
Table 8 RCTs included in the submission 44
Table 9 Manufacturer assessments of quality for vortioxetine trials 47
Table 10 Study 14178A (REVIVE). Baseline characteristics of participants 49
Table 11 REVIVE trial results: remission and response 50
Table 12 REVIVE trial: Primary and secondary efficacy outcomes: scale score differences in mean change from baseline (MMRM) 52
Table 13 REVIVE trial: summary of adverse events reported over 12-week treatment period 52
Table 14 Study TAK318. Baseline characteristics of participants 55
Table 15 Characteristics of studies included in indirect comparisons 59
Table 16 Summary of outcomes from trials included in the main network meta-analysis 61
Table 17 Summary of the results of the frequentist network meta-analyses 63
Table 18 Summary of the results of the Bayesian network meta-analyses 63
Table 19 Summary of network meta-analysis reporting standardised mean differences 64
Table 20 Overview of pools for the evaluation of the safety and tolerability of vortioxetine 66
Table 21 Summary of TEAEs by dose (APTS)- Short-term pool 67
Table 22 TEAEs with an incidence of 2% or more by preferred term, core treatment period (APTS) - Short-term pool 68
Table 23 Incidence of TEAEs of ≥5% in the open-label, long-term pool, core treatment period 70
Table 24 Summary of meta-analyses of vortioxetine vs. placebo 73
Table 25 Summary of results in the PBAC report 77
Table 26 Indirect treatment comparison by Llorca et al. 78
Table 27 NICE reference case checklist 89
Table 28 Time to 3rd-line switch (Adapted from Table B47 MS) 98
Table 29 Probability of remission and relapse assumed in the manufacturer model 102
Table 30 Percentage of patient who switch in the decision model (MS, base-case) 103
Table 31 Percentage of patients who experience treatment-emergent adverse events in the acute phase 105
Table 32 Percentage of patients who withdraw from 2nd line treatment due to AEs in the decision model (base-case, MS) 105

Table 33 Percentage of patients experiencing long-term adverse events 106
Table 34 Utility values derived from REVIVE trial 107
Table 35 Utility scores and clinical response during the study period, adapted from Sapin (2004) ... 107
Table 36 Acute phase health utilities by analysis set and method 108
Table 37 Health state utilities for both acute and other phases of the decision model 108
Table 38 Costs and disutilities associated with adverse events 109
Table 39 Acute phase drug acquisition costs 110
Table 40 Post-acute phase drug acquisition costs 110
Table 41 Health state costs, adapted from MS Table B63 111
Table 42 Model outputs by clinical outcomes, for each comparator 114
Table 43 Cost-effectiveness results from MS base-case 115
Table 44 MS scenario and structural sensitivity analyses, adapted from MS Table B83 116
Table 45 2-month time horizon scenario results 121
Table 46 8-month time horizon scenario results 121
Table 47 24-month time horizon scenario results 121
Table 48 GP visit for switch set to zero scenario results 122
Table 49 Time to 3rd-line switch (Adapted from MS Table B47) 124
Table 50 Corrections to health utilities made by ERG 128
Table 51 Cost-effectiveness results: manufacturer base-case with revised utilities 129
Table 52 Alternative scenarios regarding relative effectiveness: proportion of remitters at 8 weeks 130
Table 53 Short-term adverse events probabilities for duloxetine and escitalopram 131
Table 54 Long-term adverse events probabilities for duloxetine and escitalopram 132
Table 55 Acquisition costs for duloxetine and escitalopram 132
Table 56 Alternative scenarios regarding relative effectiveness: cost-effectiveness (with up-titration) 133
Table 57 Alternative scenarios regarding relative effectiveness: cost-effectiveness (without up-titration) 134
Table 58 Alternative ICER estimates assuming same relapse rate and average remission rate of 2nd line treatments (with up-titration) 137
Table 59 Alternative ICER estimates assuming same relapse rate and average remission rate of 2nd line with proportionate reduction based on STAR*D (with up-titration) 138

Table of Figures

Figure 1 Typical course of an MDE (source: MS Figure A3) 24
Figure 2 Interpretation of NICE CG90 guidance, with vortioxetine position highlighted..... 25
Figure 3 Overview of clinical evidence for vortioxetine in adults with MDD (from MS, Figure A2). 29
Figure 4 Modified treatment pathway from CG90 Guidance showing proposed positioning of vortioxetine..... 34
Figure 5 Pharmacological treatments prescribed for patients diagnosed with depression, by line of therapy (from MS, Figure A10)..... 35
Figure 6 Kaplan Meir plot of time to first response (50% reduction in MADRS) 51
Figure 7 Study TAK318. Changes in clinical measures of depressive symptoms over 8 weeks of treatment with vortioxetine and escitalopram..... 56
Figure 8 Base-case network used in indirect treatment comparison..... 62
Figure 9 Forest plot of relative risk of TEAEs with an incidence of 2% or more by preferred term, core treatment period (APTS) - MDD Short-term pool..... 69
Figure 10 Forest plot of standardised mean difference based on data from Pae et al. 75
Figure 11 Forest plot of response rate based on data from Pae et al..... 75
Figure 12 Forest plot of remission rate based on data from Pae et al. 76
Figure 13 Model schematic diagram (MS Figure B21) 92
Figure 14 Markov state transition diagram for modelled events following switch to third line (MS Figure B22)..... 93
Figure 15 Diagrammatic representation of the course of an MDD episode (MS Figure A3, adapted from Bakish)..... 93
Figure 16 Markov state transition diagram for modelled events following switch to third line..... 99

List of abbreviations

AE	Adverse event
AHRQ	Agency for Healthcare Research and Quality
ANCOVA	Analysis of covariance
APTS	All patients treated set
BMI	Body mass index
CGI-I	Clinical Global Impression - Improvement Scale
CGI-S	Clinical Global Impression - Severity Scale
CI	Confidence interval
CSFQ-14	Changes in Sexual Functioning Questionnaire Short-Form
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision
EMA	European Medicines Agency. Formerly abbreviated as “EMEA”.
EPAR	European Public Assessment Report
EQ-5D	EuroQoL Questionnaire 5-Dimension
FAS	Full analysis set
FDA	Food and Drug Administration
HAM-A	Hamilton Anxiety Rating Scale
$HAM\text{-}D_{17}$	Hamilton Depression Rating Scale 17-item
$HAM\text{-}D_{24}$	Hamilton Depression Rating Scale 24-item
HR	Hazard ratio
HRQoL	Health-related quality of life
ICD-10	International Classification of Diseases
ITC	Indirect treatment comparison
ITT	Intention to treat
LOCF	Last observation carried forward
LREG	Logistic regression
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	Major depressive disorder
MDE	Major depressive episode
MMRM	Mixed model for repeated measures
OR	Odds ratio
PHQ-9	Patient Health Questionnaire
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RR	Relative risk
SAE	Serious adverse event
SNRI	Serotonin–norepinephrine reuptake inhibitors
SSRI	Selective serotonin reuptake inhibitor
TEAE	Treatment-emergent adverse event
XR	Extended release

1 Summary

1.1 Critique of the decision problem in the manufacturer's submission

Vortioxetine (brand name Brintellix®) is an antidepressant with a different mechanism of action to other antidepressants such as SSRIs and SNRIs, which has been claimed to act on a number of transmitter systems. The Committee for Medicinal Products for Human Use of the EMA granted marketing authorisation on 18 December 2013 for the treatment of major depressive episodes (MDE) in people with major depressive disorder (MDD).

The manufacturer's decision problem was substantially narrower than that of the NICE scope, primarily in terms of the population considered. The patient population was restricted to a subset of the licensed patient population; namely, only patients who responded inadequately in terms of efficacy or tolerability to initial antidepressant treatment, and who switched to an alternative antidepressant. This was referred to as the "switch population". The manufacturer justified this restriction by stating that the distinct pharmacological profile and favourable tolerability profile of vortioxetine may be particularly suitable in the switch population.

The ERG accepts that vortioxetine may be used in a switch population, however it is our view that presenting evidence only for the switch population represents an important limitation from both a clinical and cost-effectiveness perspective. The ERG considers that the appropriate population and potential position of vortioxetine should have been based on a broader consideration of the evidence for vortioxetine and other comparators.

The restriction to a switch population severely constrained the evidence presented. The submission presented only two trials directly comparing vortioxetine to other antidepressants. These two trials represent only 972 patients of over 7,000 patients included in studies of vortioxetine. Only four trials were included in the primary indirect comparison of treatments.

The final scope issued by NICE identified a wide range of relevant comparators including SSRIs, SNRIs, tri-cyclic antidepressants, other types of antidepressant and augmentation treatments. The restriction in scope to a switch population also meant that comparators were restricted primarily to those most likely to be used as second-line therapies, based on NICE guidelines, including SSRIs, SNRIs and newer-generation antidepressants. The restriction to consider only switch population evidence, and the small number of trials identified, meant that vortioxetine was compared directly only to agomelatine and escitalopram; and indirectly only to agomelatine, sertraline, venlafaxine (XR), bupropion and citalopram. In particular, no comparisons to duloxetine, fluoxetine or mirtazapine were made.

The final NICE scope included a range of relevant outcomes, including response to treatment, remission, relapse, symptoms severity, anxiety, cognitive dysfunction, sleep quality, health-related quality of life and a range of adverse events. The manufacturer reported relevant data for most of these outcomes in the two trials of vortioxetine, but the review of indirect evidence included only two outcomes (remission and withdrawal due to adverse events). This further limited the evidence on the relative efficacy, safety and cost-effectiveness of vortioxetine versus other active comparators.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The manufacturer's submission on clinical effectiveness included four systematic reviews: a review of RCTs of vortioxetine compared to active comparators in the switch population to evaluate efficacy; a review of non-RCT evidence of vortioxetine in the switch population; a review of adverse events of vortioxetine; and a review of indirect comparative evidence, again in the switch population, including an indirect treatment comparison and network meta-analysis of RCTs of other antidepressants, to evaluate efficacy and safety.

In response to the points for clarification, the manufacturer also provided meta-analyses of trials in the non-switch population.

1.2.1 Direct evidence in the switch population

As systematic review of RCTs of vortioxetine compared to active comparators in the switch population to evaluate efficacy was reported. It identified two trials, REVIVE and TAK318, summarised in Table 1. These trials were not combined in a meta-analysis because they were in different populations and used different comparators.

Table 1 RCTs included in the submission

Study	Regimen & duration	Comparator	Design	Follow-up duration	Primary outcome	Patient population
REVIVE (14178A)	Vortioxetine 10mg-20mg flexible dosing, 12 weeks	Agomelatine 25-50mg flexible dosing	Double-blind, international phase IIIb, parallel-group randomised trial	Efficacy: up to 12 weeks from baseline Safety: up to 16 weeks from baseline	Change from baseline in depression severity (MADRS total score) at week 8	Patients who have experienced an inadequate response to an SSRI or SNRI in their current MDE
TAK318	Vortioxetine 10mg-20mg flexible dosing, 8 weeks	Escitalopram 10-20mg flexible dosing	Double-blind, multicentre phase IIIb, parallel-group randomised trial	Efficacy: up to 8 weeks from baseline Safety: up to 12 weeks from baseline	Change from baseline in sexual functioning (CSFQ-14 total score) after 8 weeks of treatment	Patients who are well-controlled on an SSRI but experienced treatment emergent sexual dysfunction

MADRS: Montgomery and Åsberg Depression Rating Scale; CSFQ-14: Changes in Sexual Functioning Questionnaire Short-Form

The REVIVE trial was a well conducted trial comparing vortioxetine with agomelatine. The primary outcome in REVIVE was change in depression scores. Vortioxetine showed statistically significant superiority to agomelatine, reducing symptoms of depression. Average MADRS scores were 2.16 points lower on vortioxetine (95% CI 0.81 to 3.51) than on agomelatine after eight weeks. Vortioxetine was also superior to agomelatine in terms of response rate using MADRS at 8 weeks (OR 1.81, 95% CI 1.26 to 2.60) and remission rate (OR 1.72, 95% CI 1.17 to 2.52).

Vortioxetine and agomelatine had similar rates of treatment-emergent adverse events (around 54%) and serious adverse events (around 1.5%), but vortioxetine had lower rates of adverse events leading to withdrawal (5.9% vs 9.5%).

The TAK318 trial compared vortioxetine with escitalopram. The primary outcome was sexual functioning measured using the CSFQ-14 scale. At 8 week treatment the mean change from baseline was 8.8 (SE 0.64) in the vortioxetine arm and 6.6 (SE 0.64) in the escitalopram arm. This difference was statistically significant in favour of vortioxetine ($p = 0.013$). The submission also reported difference between arms on the CSFQ-14 subscales; there was statistically significant evidence in favour of vortioxetine on all subscales presented.

No differences between vortioxetine and escitalopram were identified for changes in depression scales, response, relapse or remission rates. Adverse event rates were similar on both treatments.

A systematic review for non-RCT evidence was performed, but no relevant studies were identified.

A systematic review for safety trials in the switch population was performed but no trials were identified. The search was expanded to the general, non-switch population and five open-label extensions of vortioxetine trials were identified. On clarification the manufacturers also supplied adverse event data from 12 short-term placebo controlled trials of vortioxetine.

About 6% of patients had severe adverse events on vortioxetine compared to 4% on placebo in the placebo controlled trials. In the open-label extension trials 8.1% of patients had a serious adverse event. The manufacturer concluded that vortioxetine had a generally good safety profile. No data were submitted comparing adverse events using vortioxetine with other active treatments.

1.2.2 Indirect evidence in the switch population

The submission included a systematic review of treatments other than vortioxetine in the switch population used to perform an indirect treatment comparison with vortioxetine for efficacy and safety. This review identified seven trials. One was excluded on the grounds of poor trial quality. Two were placebo controlled trials not included in the primary network meta-analysis. Four trials were used in the primary network meta-analysis; there was only one trial included for each treatment comparison.

The outcomes of the network meta-analysis were remission rate and withdrawal rate due to adverse events. A range of models were fitted using different assumptions, including the placebo controlled trials and using both frequentist and Bayesian statistical methods. A summary of the results from the main frequentist analysis are shown in Table 2. Results from other analyses were broadly consistent with these results.

Table 2 Summary of the results of the frequentist network meta-analyses

	Remission rate			Withdrawal rate due to adverse events		
	Rate (%)	**Risk Difference vs vortioxetine (%)**	**95% CI**	**Rate (%)**	**Risk Difference vs vortioxetine (%)**	**95% CI**
Vortioxetine	40.5	–	–	5.9	–	–
Agomelatine	29.5	-11	-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

Vortioxetine had higher rates of remission than all other treatments, but results were only statistically significant for agomelatine. Vortioxetine also had lower rates of withdrawal due to adverse events than all other treatments, but results were only statistically significant for comparisons with sertraline, venlafaxine (XR) and bupropion. Due to the limited number of trials in the network, no assessments of heterogeneity or network inconsistency were performed.

1.2.3 Evidence syntheses of non-switch populations

As discussed in Section 1.1 the ERG questions the validity of restricting the analysis to switch populations. The ERG therefore requested that the manufacturer provide results from trials or meta-analyses of trials comparing vortioxetine to other active treatments and/or placebo in initial use and non-switch use populations. The manufacturer provided meta-analyses of their trials comparing

vortioxetine to placebo, but not for vortioxetine versus active comparators. The manufacturer also reported the existence of two submissions to regulators, four systematic reviews of vortioxetine trials and one indirect comparison of vortioxetine with other antidepressants.

The meta-analyses of trials submitted by the manufacturer and the four identified systematic reviews compared vortioxetine with placebo. There was considerable overlap in the trials included in the meta-analyses, and they reported different outcomes. However, all concluded that vortioxetine was superior to placebo.

Two systematic reviews and one regulatory submission (to the Pharmaceutical Benefits Advisory Committee of Australia) compared vortioxetine to other antidepressants by using active reference arms from placebo controlled trials. Active reference arms are included in trials of antidepressants to ensure that patients are responding to therapy. An active reference should be a drug of proven superiority over placebo, so it can be used to check whether the trial has successfully treated patients by confirming a difference between the active reference and placebo. After reanalysis by the ERG of data from one of these reviews (Pae et al. 2014) vortioxetine was found to be inferior to duloxetine in terms of changes in depression scores, response rate and remission rate. There was no evidence of a difference between vortioxetine and venlafaxine. Results from the other review and the regulatory submission were consistent with these results.

One systematic review (Llorca et al. 2015) performed an indirect comparison of vortioxetine with other antidepressants by analysing all placebo controlled trials of vortioxetine and other antidepressants. The results of the indirect treatment comparison are summarised in Table 3. This analysis found no statistically significant evidence of any difference in efficacy between vortioxetine and other antidepressants (except for agomelatine), but did find evidence that vortioxetine had a lower withdrawal rate due to adverse events than venlafaxine, desvenlafaxine and sertraline.

Table 3 Indirect treatment comparison by Llorca et al

	Results vs vortioxetine (Standard error)						
	Agomelatine	**Desvenlafaxine**	**Duloxetine**	**Escitalopram**	**Sertraline**	**Venlafaxine**	**Vilazodone**
Efficacy at 2 months (SMD: <0 favours vortioxetine)	-0.156 (0.113)	0.025 (0.803)	0.090 (0.419)	-0.054 (0.695)	-0.037 (0.832)	0.124 (0.328)	-0.245 (0.111)
Withdrawal (Odds ratio: <1 favours vortioxetine)	1.769 (0.030) *	0.578 (0.035) *	0.752 (0.262)	0.671 (0.275)	0.299 (0.008) **	0.469 (0.009) **	0.640 (0.181)

Response rate (Odds ratio: >1 favours vortioxetine)	1.045 (0.815)	1.153 (0.364)	0.893 (0.514)	0.843 (0.523)	0.772 (0.575)	0.789 (0.353)	0.975 (0.934)
Remission rate (Odds ratio: >1 favours vortioxetine)	1.220 (0.470) **	1.029 (0.852)	0.894 (0.526)	0.990 (0.981)	NA	0.689 (0.444)	0.983 (0.952)

** p-value 0.01 – 0.05; ** p-value <0.01*

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The restriction to a switch population meant that only two trials comparing the efficacy of vortioxetine with other antidepressants were submitted. The REVIVE trial comparing vortioxetine to agomelatine found vortioxetine had greater short-term reduction in depression scores and lower withdrawal rates due to adverse events, but this result may not be reliable as this was a non-inferiority trial and was not powered to detect superiority of vortioxetine. It should also be noted that agomelatine is not approved for use in the UK as no evidence for its efficacy has been submitted. The TAK318 trial included patients who were switching due to sexual dysfunction but had responded to initial treatment. The ERG notes that this is a very narrow and specific population, so the TAK318 trial provided little information on the broader population who might take vortioxetine. It did find that vortioxetine reduced sexual dysfunction symptoms when compared to escitalopram. However it is not clear whether this finding is specific to vortioxetine, or whether any non-SSRI would have a similar beneficial effect. No efficacy evidence was submitted for the licensed 5mg vortioxetine dose.

Based on the safety evidence presented, vortioxetine appears generally safe and tolerable in patients with MDD. Most adverse were mild to moderate in intensity and there was no conclusive evidence that these were dose dependent. The submission did not present any safety comparisons of vortioxetine with any active comparators, so the safety profile of vortioxetine compared to other antidepressants is uncertain.

The ERG has considerable concerns over the validity of the network analysis because of the high apparent diversity in the populations across trials, with very different included patients and severities of depression. For one trial (Kasper) the analysis used a subset of patients who had been treated in the year prior to enrolment. This is not the same as patients who were switching treatment, so the ERG does not think this trial should have been included. The ERG questions the validity of the indirect treatment comparison, but notes that the analysis found no convincing evidence of difference between vortioxetine and other treatments in terms of remission rate (except for agomelatine). There was some

evidence that vortioxetine may have lower withdrawal rates due to adverse events, but the high apparent heterogeneity across trials means the validity of this finding is questionable.

Given the limited nature of the data in the switch population the ERG considers that data in non-switching and initial-use populations should be considered. Although such data is not in the switch population it is relevant to the broader population of all patients with MDD specified in the NICE scope. The manufacturer justified excluding trials of non-switching populations by claiming that treatment efficacy in a switch population may be different from in initial use. The ERG considers that the evidence submitted to justify this claim is limited and refers only to patients who had previously used an SSRI, where switching to another SSRI may be less effective than a non-SSRI treatment. No evidence was presented to suggest that the relative efficacy of non-SSRIs may vary between initial and switch use, and no evidence was specific to vortioxetine. The ERG therefore concludes that this restriction was inappropriate and evidence on non-switch populations is relevant when examining the efficacy and safety of vortioxetine.

Direct evidence comparing vortioxetine to other active treatments in non-switching populations was only available from short-term placebo controlled trials with active reference arms. The manufacturer has criticised the use of active references because they are not true randomised comparisons and patients known to be non-responsive to the reference are excluded, possibly biasing results in favour of the active reference. While the ERG accepts the potential for such bias it does not consider this potential bias to be substantial enough to exclude these trials. The ERG found no evidence of any difference in efficacy between vortioxetine and venlafaxine, based on two trials. There was evidence that vortioxetine was significantly inferior to duloxetine in terms of reducing depression scores, response and remission. While there is a possibility of bias in favour of duloxetine in these analyses it is not clear whether any bias would be sufficient to completely explain this inferiority.

An indirect treatment comparison in non-switch populations found no evidence of any difference in efficacy between vortioxetine and other treatments, but there was some evidence to suggest that vortioxetine had a lower withdrawal rate due to adverse events than some treatments, including sertraline and venlafaxine. While this is an indirect analysis, and not conducted in a switch population, the number of trials included in this analysis suggests that this may represent the most reliable evidence for comparing the efficacy and tolerability of vortioxetine to other treatments.

The ERG concludes, based on the totality of the evidence, that vortioxetine is likely to be of similar efficacy to other antidepressants, but may be superior to agomelatine and inferior to duloxetine. Vortioxetine appears to have a lower withdrawal rate due to adverse events than other treatments, and so may be more tolerable, however data on adverse events with vortioxetine, particularly when

compared to other antidepressants, are too limited to draw any firm conclusions on the safety of vortioxetine.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The MS presented evidence on the cost-effectiveness of vortioxetine in a primary care setting using only the following comparators: agomelatine, sertraline, citalopram and venlafaxine XR. Other relevant comparators, such as escitalopram or duloxetine, were excluded due to the absence of evidence in the switch population.

To evaluate cost-effectiveness, the MS presented a decision model that evaluated the progression of a single MDE. The model was based on treatment success defined in terms of remission at 8-weeks. It followed up patients for 12 months and considered three stages of disease progression: the acute phase (2 months duration), a maintenance phase (6 months duration), and a recovery phase (4 months duration). The model used a decision-tree to evaluate progression within second-line of treatment, and a separate Markov process to describe further lines of therapy that may subsequently be used.

The initial decision tree-structure was common to all patients during the initial acute phase period (0-2 months). During this period patients may achieve remission or they may withdraw from their current therapy due to short-term side-effects or failure to achieve remission. Patients achieving remission in the initial acute phase period subsequently continue within the main decision-tree structure. Between months 2 and 8, these patients subsequently follow additional pathways (or branches) covering the maintenance phase. During the maintenance phase, patients are assumed to stay on treatment (i.e. sustained remission) or they may stop treatment due to an adverse event or subsequent relapse. If patients sustain remission during the maintenance period, they enter the final part of the decision tree structure representing the recovery phase which covers the final 4 months of the total 12-month time horizon. Importantly, during the recovery period, the therapy is assumed to be discontinued and an assumption is made that patients are no longer at risk of relapse or recurrence.

The manufacturer's model assumed that a proportion of patients withdraw from treatment due to adverse events. In the acute phase, these patients were assumed to be non-remitters. The timing of many events within the decision tree was static – for example, all patients were assumed to withdraw one month into the maintenance phase and relapse three months into this phase. The final phase of the decision tree in the MS is the recovery phase. During this phase patients were no longer assumed to be treated with antidepressants. An important additional assumption was also made that patients no longer face a risk of relapse or incur any other NHS or PSS costs.

The separate Markov process was used to model subsequent lines of therapy, with the model allowing additional lines of treatment (i.e. covering 3rd, 4th and 5th lines line of therapy, since the initial therapy in the switching population is already the 2nd line of treatment in the overall management pathway). The effectiveness of further lines of therapy (3rd, 4th and 5th lines) is independent of the initial treatment strategy. The Markov model uses a 2-month cycle and each time a patient moves to the no remission state (including movements from the same state) they are switched to a further line of treatment.

1.4.1 Evidence used to inform the decision model

The main source of evidence on vortioxetine for the acute phase of treatment in the MS was the REVIVE trial. Because this only provided direct evidence for vortioxetine compared to agomelatine, the MS used an indirect comparison to infer remission and withdrawal rates against other comparators. The probability of relapse for the maintenance phase in 2nd line treatment was assumed to be the same for vortioxetine and all comparators, taken from Limosin (2004).

In the decision model, the probability of withdrawing from treatment in the acute phase due to AEs is informed by the indirect comparison. All patients that withdraw were assumed to be non-remitters. A different source of evidence was used to quantify the incidence of specific adverse events (sexual dysfunction, dry mouth, nausea, sweating, somnolence, headache, diarrhoea, insomnia and dizziness), so that associated costs and utility decrements could be evaluated. For agomelatine and vortioxetine REVIVE was used and for other treatments Cipriani (2010), Cipriani (2012) and pooled Cochrane reviews were used. Long-term adverse events considered were sexual dysfunction, insomnia and weight gain. Their incidence was informed by pooled long-term extension studies. It is assumed that if a patient has an adverse event then they will switch treatment with a probability of 25%.

Remission and relapse with further lines of treatment were informed by the blend of treatments used in STAR*D. This study reported that the probabilities of achieving remission and sustaining remission appear to decline in later lines of treatment, and within the MS these were assumed to be independent of the initial switch treatment received.

The MS considered different sources of evidence for utilities associated with short term health outcomes (up to 8 weeks), where the REVIVE trial was used, and longer term outcomes (after 8 weeks), where Sapin (2004) was used. The decision model considers a number of different resource use categories. Acquisition costs of drugs are taken from standard sources and applied to the dose of the drug (licensed dose in the acute phase and up-titration in the maintenance phase). Other cost categories considered were health state costs related to an MDE, which include GP, psychotherapist and psychiatrist consultations, in addition to hospitalisations.

Brief description of cost effectiveness results in the MS

The cost effectiveness results from the MS base-case are shown in Table 4. These results demonstrate that vortioxetine is both cheaper and more effective than agomelatine. Outcomes are driven by the ability to improve remission and the better adverse event profile assumed for vortioxetine.

Table 4 MS base-case results

	Venlafaxine	Vortioxetine	Citalopram	Sertraline	Agomelatine
Cost effectiveness					
Expected QALYs per patient	0.675	0.694	0.664	0.664	0.676
Expected costs per patient	£964	£971	£976	£977	£1,082
ICER	reference	£378	Dominated	Dominated	Dominated

The manufacturer also conducted a number of sensitivity analyses. The MS base case model was robust to the range of scenarios investigated.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The MS evaluated cost effectiveness for the switch population, and did not consider more fully the broader population and potential position of vortioxetine within current pathways. The MS included only relative effectiveness evidence on the switch population. Given the ERG deems the indirect comparison invalid, cost effectiveness evidence submitted (summarised in Table 4) is only interpretable for the comparison of vortioxetine and agomelatine. The ERG believes a broader evidence-base (including initial use) should have been considered to overcome these limitations.

The ERG has a number of additional significant concerns regarding the model structure employed by the manufacturer. Most importantly, the manufacturer's model does not explicitly consider response to treatment and instead only considers remission at 8 weeks. This does not seem to reflect clinically appropriate definitions of initial treatment success and subsequent clinical decisions.

As to the evidence used to describe further lines of treatment, the STAR*D study, the ERG considers that it includes treatments with limited overlap and that the patient population may not be generalisable. Also, the STAR*D shows a lower probability of achieving remission and a higher relapse rate for subsequent switches than expected in clinical practice. The ERG considers the use of separate sources of HRQoL data for particular inputs in the model appeared inconsistent.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

The submitted evidence examined the effectiveness of several relevant antidepressants compared to vortioxetine in the switch population. The submission covered the key clinical outcomes, including changes in depression scores, remission rate, withdrawal rate and incidence of adverse effects. Appropriate statistical methods were used to perform a network meta-analysis and suitable sensitivity analyses were performed.

The ERG considers the manufacturer presents cost effectiveness evidence for vortioxetine for a restrictive case regarding second line use in a primary care setting. The decision model used by the manufacturer seems to broadly reflect the progression of an MDE, and the ERG considers that the use of 12-month horizon appears reasonable for the 'average' patient given that the average duration of an untreated MDE is considered to be between 5-6 months.

1.7 Weaknesses and areas of uncertainty

The ERGs primary concern with the evidence submitted is related to the manufacturer's decision to restrict the submission to a switch population, rather than all patients with MDD as specified in the scope. This restriction meant that only two trials comparing vortioxetine to other antidepressants were presented. One compared vortioxetine to a drug not licenced in the UK, and the other was focused only on patients with SSRI treatment-emergent sexual dysfunction. Therefore neither appeared representative of most UK patients likely to be switching antidepressant treatments. The trials included in the network meta-analysis were very diverse in both the populations included and in their results. The ERG concludes that, as a result, this analysis did not present reliable evidence on the relative efficacy of vortioxetine compared to other antidepressants.

The manufacturer assumed that the most appropriate position of vortioxetine is as a second-line treatment. The ERG's view is that the manufacturer should have developed a more flexible model which was capable of assessing the value of vortioxetine in alternative positions within current treatment pathways. The manufacturer based their economic analysis on relative effectiveness evidence specifically related to the switch population. The ERG, however, feels that the broader evidence base on initial use of vortioxetine and comparator treatments should have also been considered.

The ERG considers there are significant uncertainties concerningboth the decision rule applied and the assumption that only patients who achieve remission at 8 weeks will be continued on their initial therapy, and the use of 8-week data to inform switching decisions at an earlier time point. The ERG suggests that STAR*D may impose a worse prognosis to further lines of therapy (i.e. a lower remission rate and higher relapse rate) than what might be expected for the population of interest.

Also, the ERG considers the use of separate sources of HRQoL data for particular inputs appeared inconsistent and potentially optimistic towards the cost-effectiveness of vortioxetine.

1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG explored some of the key issues and uncertainties on the manufacturer's cost-effectiveness analysis.

First, the ERG revised the manufacturer base-case to correct issues found with the utility scores, namely the use of a single source of evidence (the REVIVE study). The proposed corrections appear to have only minimal effect on the QALY estimates. The second analyses undertaken by the ERG was to consider the broader evidence base on the relative effectiveness of vortioxetine and comparator treatments using three scenarios: scenario 1 uses placebo controlled trials (Llorca et al.); scenario 2 uses direct evidence (Pae et al.); and scenario 3 assumes equal effectiveness. Table 5 reports the results of these analyses.

Table 5 Alternative scenarios regarding relative effectiveness: cost-effectiveness (with up-titration)

			Incremental		ICER	
	Costs	**QALYs**	**Costs**	**QALY**	**w SSRI**	**w/o SSRI**
			(in relation to *ref*)		(incremental analyses, in relation to next best)	
Scenario 1: Llorca						
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
Scenario 2: Pae						
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
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

The results are clearly sensitive to the assumptions made concerning the relative effectiveness of the alternative treatments in achieving remission. In Scenario 1, venlafaxine (XR) dominates vortioxetine

and the other comparator treatments: the higher remission probabilities assumed for venlafaxine (XR) appear to more than offset any additional benefits attributed to vortioxetine and/or other comparator treatments in terms of adverse events. In Scenario 2, venlafaxine (XR) remains the lowest cost strategy but no longer dominates vortioxetine and duloxetine. The most cost-effective treatment in Scenario 2 is now duloxetine which has an ICER of £13,393 compared to vortioxetine. In Scenario 3, when all treatment are assumed to be equally effective in terms of achieving remission at 8-weeks, the differences are now driven entirely by the different acquisition costs and the assumptions related to adverse events. Within this scenario escitalopram is now the lowest cost strategy and appears to dominate venlafaxine (XR). The ICER of vortioxetine is £18,888 per QALY compared to escitalopram. When escitalopram is excluded from consideration, venlafaxine (XR) is now the cheapest strategy and the ICER of vortioxetine is £5,318 per QALY compared to venlafaxine (XR). These results are, however, sensitive to assumptions related to the effectiveness of subsequent lines of therapies tested by the ERG.

1.9 Conclusions from ERG analyses

The ERG's exploratory analyses have demonstrated the cost-effectiveness results for vortioxetine presented by the manufacturer are contingent on the assumption that vortioxetine is the most effective initial switch treatment in terms of remission. Given the higher acquisition cost of vortioxetine relative to other SSRI therapies and venlafaxine (XR), the cost-effectiveness of vortioxetine was only clearly evident in those scenarios where it was assumed that vortioxetine was more or equally effective in achieving remission. However, when the effectiveness estimates were based on a broader set of trials than those considered by the manufacturer, the cost-effectiveness of vortioxetine was less evident. Across all of scenarios tested by the ERG, differences in terms of tolerability have been assumed, i.e. additional benefits with vortioxetine from improved tolerability and/or reductions in adverse events (short and longer-term).

2 Background

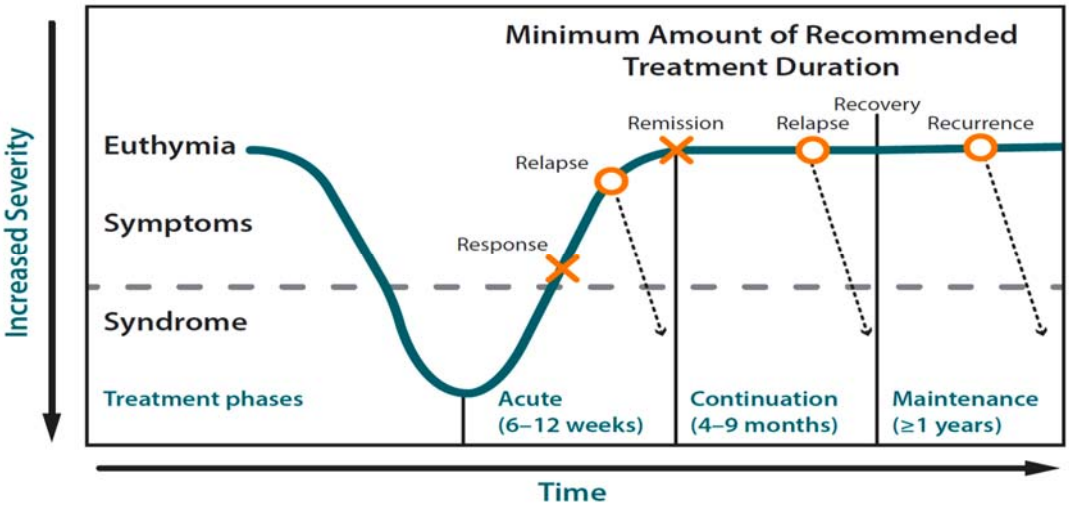
2.1 Critique of manufacturer's description of underlying health problem.

The manufacturer presents a suitable summary of major depressive disorder (MDD) and major depressive episodes (MDE) in manufacturer's submission (MS) sections 2.1 to 2.3; this includes their definition, progression over time, impact on health and quality of life and the impact and cost to the NHS and society in general.

The submission recognises that the terms MDD and MDE are based on standard classifications from the Diagnostic and Statistical Manual of Mental Disorders (DSM); these are used to describe depression in the USA where terminology may differ slightly from that used in the UK, where the terms mild/moderate/severe depression are more commonly used. The submission notes that MDD/MDE are approximately equivalent to moderate to severe depression as defined by NICE guidance.(1) An MDE is defined as the occurrence of depressed mood or a loss of interest or pleasure in life activities for at least two weeks, at least five of the nine core symptoms of depression and at least one significant impairment in social functioning (list provided in submission) occurring almost every day. Left untreated an MDE may last for two months to several years, but this may be greatly reduced by treatment.(2) The course of an MDE is characterised by five events: response, remission, recovery, relapse and recurrence,(3) as summarised in Figure 1. A patient whose symptoms improve sufficiently on treatment is considered to be responding, and is judged to be in remission when core symptoms of depression cease or are considered very minor. Patients remaining in remission for sufficient time are judged to have recovered. NICE recommends continuing treatment for at least six months following remission, and treatment may be continued for longer (two years or more) in cases of recurrent MDEs.(1) A patient may relapse at any time if symptoms worsen. If symptoms re-occur after recovery this is judged to be a recurrence and constitutes a new MDE. Relapse and lack of response are common, with approximately one-third to one half of patients not responding adequately to treatment,(4, 5) and at least half of all patients experiencing their first MDE will go on to have at least one more episode within the next ten years.(6)

The submission discusses the potential health impacts of depression, particularly the substantial loss of quality of life and potential reduction in life expectancy, including increased risk of suicide, and its impact on family members. The costs associated with depression are substantial, with 1.24 million people estimated to have depression in England, costing the NHS £1.68 billion, and averaging £2,805 per person.(7) The submission also noted the substantial wider economic cost of depression due to increased sick leave and absenteeism from work, and reduced productivity.

Figure 1 Typical course of an MDE (source: MS Figure A3)

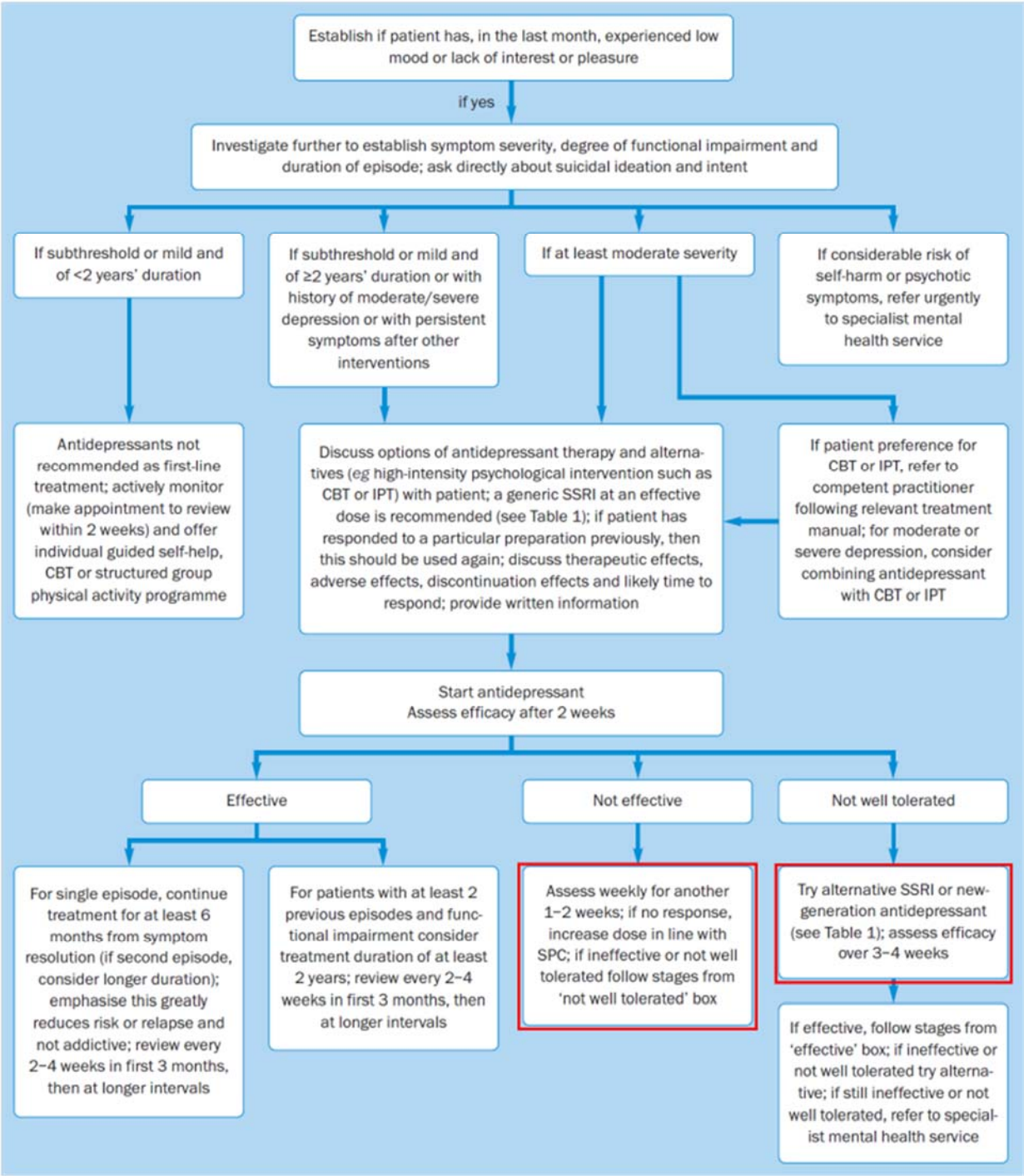


2.2 Critique of manufacturer's overview of current service provision

The submission presents a generally suitable summary in MS sections 2.5 and 2.6 of current clinical practice and service provision with reference to the most recent NICE guidelines, specifically Clinical Guidance 90 (CG90, 2010).(1)

The current NICE guidelines (CG90) are summarised in Figure 2**Error! Reference source not found.** (taken from MS, figure A5). These recommend that patients with moderate or severe depression, but without psychotic symptoms or risk of self-harm, should be offered an SSRI and/or a high-intensity psychological intervention (for example, cognitive behavioural therapy or interpersonal therapy). If the SSRI is not tolerated or has been judged ineffective after at least four weeks of use then patients may switch to another SSRI or to another new-generation antidepressant, or be given the option of increasing their dose if well tolerated and if there is some evidence of initial response. If this second-line treatment is ineffective a third may be tried and/or patients should be referred to a specialist mental health service. As vortioxetine is not an SSRI, the manufacturers have proposed that vortioxetine be considered for this second-line therapy where initial SSRI treatment is unsuccessful (see highlighted section of Figure 2).

Figure 2 Interpretation of NICE CG90 guidance, with vortioxetine position highlighted



Rates of response and remission with antidepressants are low with around 50% response to initial SSRI treatment after eight weeks.(5) Remission rates decline for second-line, third-line and subsequent treatments. Antidepressants are also associated with high levels of adverse effects, with an estimated 16% of patients on SSRIs experiencing intolerable side effects. Consequently treatment

adherence is poor, with high rates of discontinuation. Many patients also switch treatments due to poor efficacy or side effects, with around 15% of patients on first-line SSRIs switching treatments.(8)

There are currently a wide variety of antidepressant drugs available with a variety of modes of action and varying efficacy and side-effect profiles. Given the low response rates and high rates of side effects, the ERG agrees with the manufacturer's opinion that there is a need for new treatments with favourable efficacy and side-effect profiles for patients who fail initial SSRI therapy.

3 Critique of manufacturer's definition of decision problem

3.1 Population

The population described in the final NICE scope is adults with major depressive disorder, reflecting the licensed population. Vortioxetine has a marketing authorisation in the UK for the treatment of major depressive episodes in adults. However, the patient population considered by the manufacturer was restricted to a subset of the licensed patient population, namely:

"*adult patients with moderate-to-severe MDD who are experiencing an MDE, who have responded inadequately in terms of efficacy or tolerability to initial antidepressant treatment, and who require and want to switch to alternative antidepressant*" (MS, p14).

The manufacturer cites the burden of initial treatment failure on patients, the health service and wider society to highlight the potential unmet clinical and economic need for more effective and better-tolerated options for patients requiring a switch of treatment where initial antidepressant treatment has failed. The manufacturer further justifies this restriction, stating that the distinct pharmacological profile and favourable tolerability profile of vortioxetine may be particularly suitable in the "switch population". This switch population is further defined by the manufacturer (MS, p38) as:

- *Patients who are experiencing a moderate-to-severe MDE, and*
- *whose current episode has been treated initially with an SSRI or SNRI, and*
- *who are candidates for a switch in the clinician's opinion, and*
- *who wish to change antidepressant treatment because of inadequate response or intolerability to the initial treatment.*

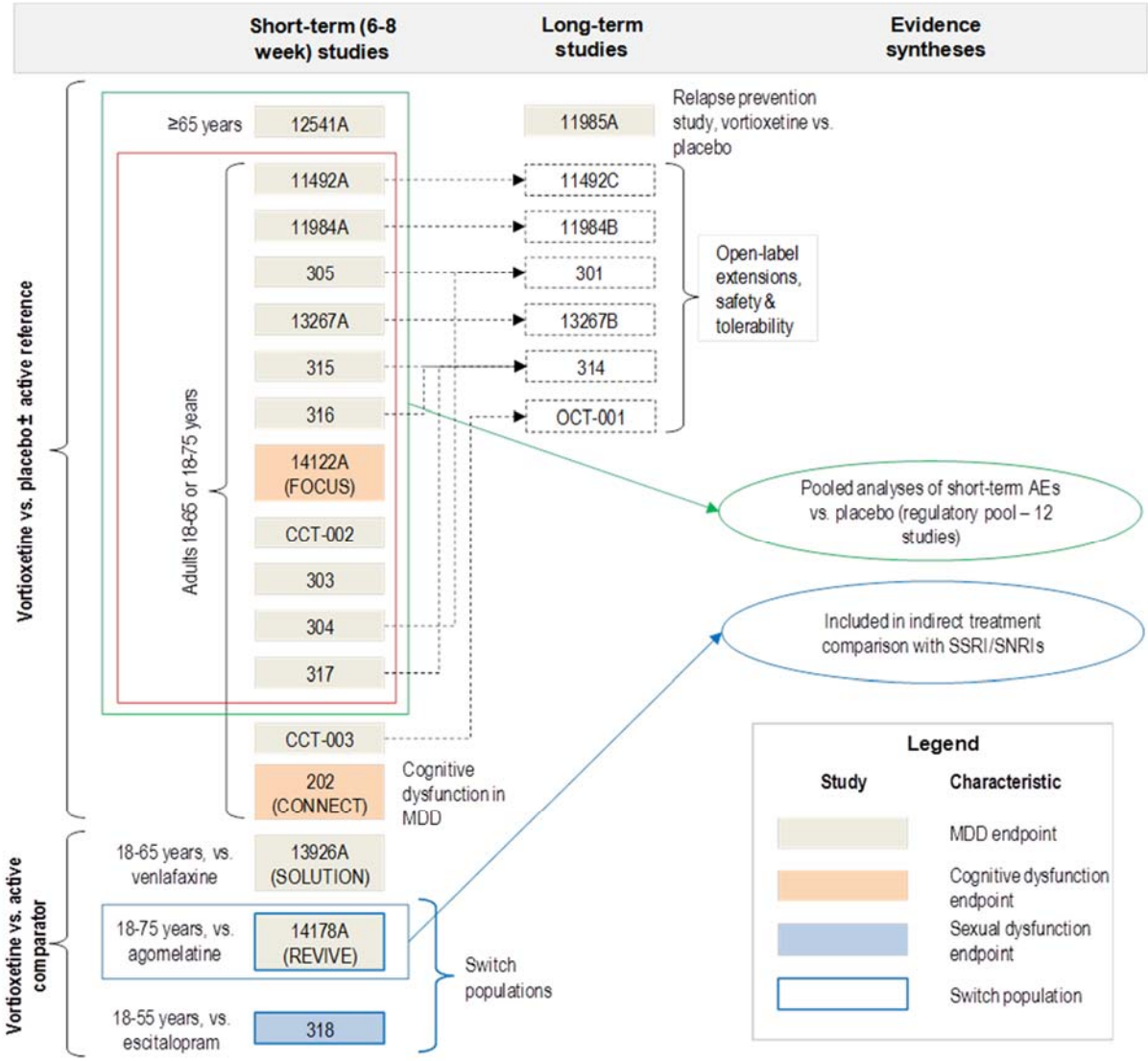
Although vortioxetine has been studied in 24 completed trials involving over 7,000 patients, the manufacturer's submission focuses largely on 2 studies considered relevant to the decision problem; Study 14178A (REVIVE) and TAK318. These studies are subsequently used as the basis for short-term efficacy and tolerability data. Longer-term safety studies from a broader population including non-switch patients are also presented in the manufacturer's submission based on the assumption that tolerability and safety data are likely to be generalisable to a switch population. Based on clinical advice, the ERG considers this assumption to be appropriate for safety, but not for tolerability, as patients intolerant of one medication may be more intolerant of others.

The population considered within the manufacturer's submission is thus significantly restricted compared to the broader population stated in the final NICE scope and the licensed indication for vortioxetine. However, such a restriction is potentially consistent with the "Other considerations" section specified in the NICE scope which states that: *"If evidence allows the subgroup of people with*

moderate or severe major depressive disorder will be considered" and *"If evidence allows the clinical and cost effectiveness of vortioxetine may be considered in different positions in the treatment pathway."*

Although the ERG acknowledges the justification provided by the manufacturer for restricting the patient population, it is our view that this represents an important limitation from both a clinical and cost-effectiveness perspective. The ERG considers that the appropriate population and potential position of vortioxetine should have been more formally demonstrated by the manufacturer, based on consideration of the full evidence base for vortioxetine and other comparators, rather than restricting the decision population and evidence base from the outset. Figure 3 provides an overview of the existing clinical evidence for vortioxetine that clearly highlights the select evidence base which subsequently underpins the manufacturer's submission.

Figure 3 Overview of clinical evidence for vortioxetine in adults with MDD (from MS, Figure A2)



The manufacturer assumed that that the MDE/MDD population considered within their submission is consistent with the terminology "moderate-to-severe depression" used in NICE CG90. This was justified on the basis that the patients included in the short-term phase III studies of vortioxetine had moderate-to-severe MDD, and that patients in the REVIVE study had a mean MADRS total score at study entry of 29 points (ranging from 22 to 43 points), indicating moderate-to-severe depression. The manufacturer detailed subsequent responses to the EMA based on the assertion in the CHMP that fully responsive patients could have been included in the study, since the severity of depression was not assessed prospectively at the onset of the first monotherapy during the lead-in period. The manufacturer reported that they considered that the absence of a baseline score prior to the initial treatment phase did not invalidate the population definition. They further stated that:

"The requirement of a baseline MADRS total score at study entry ≥22 implies that patients fully responsive to previous treatment would need to have a total score ≥44 prior to the initiation of the previous treatment. Patients with a MADRS total score ≥44 points are very rare: only 0.6% of the patients included in all the short term placebo-controlled studies in the clinical programme in MDD with vortioxetine had such a high MADRS total score" (MS, p23).

3.2 Intervention

The marketing authorisation for vortioxetine licenses it for the treatment of major depressive episodes (MDE) in adults. The marketing authorisation only covers MDEs that are associated with major depressive disorder (MDD), but not with other distinct indications such as bipolar disorder. As previously highlighted, the manufacturer's decision problem addresses a subset of the population included in the marketing authorisation and the NICE scope.

The CHMP has recently adopted a positive opinion for a Type-II variation related to the update of the SmPC for vortioxetine. The update of the SmPC provides new data on vortioxetine related to its effect on certain aspects of cognitive function and patient functioning. The application was based primarily on data from the recently completed CONNECT trial (referred to in MS, Sections 1.4 and 4), in addition to four clinical studies that were previously submitted as part of the original approval process, as well as a newly completed clinical pharmacology functional magnetic resonance imaging (fMRI) study in remitted patients with depression. The ERG does not consider that this variation has any significant implications for the stated decision problem.

Table 6 provides a summary provided by the manufacturer (MS, Table A1, p.28) regarding the formulation, cost, method of administration, posology and information of length of course (including repeat), dose adjustment and use in special populations. The manufacturer also reported that the EMA has accepted a risk management plan, which includes a non-interventional post-authorisation safety study (PASS) of vortioxetine in Europe.

Table 6 Summary of vortioxetine

Pharmaceutical formulation	Film-coated tablets, 5mg, 10mg, 20mg. (Also approved but not available: 15mg tablets).
Pack sizes	5mg x 28 10mg x 28 20mg x 28
Acquisition cost (excluding VAT)	5mg x 28: £27.72 per pack 10mg x 28: £27.72 per pack 20mg x 28: £27.72 per pack
Method of administration	Oral
Posology	The starting and recommended dose is 10mg once daily in adults less than 65 years of age. The lowest effective dose of 5mg vortioxetine once daily should always be used as the starting dose in patients ≥65 years of age.
Average length of a course of treatment	After the depressive symptoms resolve, treatment for at least 6 months is recommended for consolidation of the antidepressive response (see SmPC).
Average cost of a course of treatment	Approximately £220 for patients who achieve remission after an acute phase of treatment, assuming treatment is continued for 240 days or 8 months (assumed 8 weeks in acute phase and 6 months' consolidation) in total.
Expected average interval between courses of treatments	According to need, depending on whether remission is sustained through the recommended period of maintenance treatment or whether an MDE recurs. (See section 2.1 for additional information on the course of disease).
Expected number of repeat courses of treatments	According to need, depending on recurrence of MDEs. Recurrence of episodes is a feature of MDD, but the course of disease is highly variable (see section 2.1).
Dose adjustments	Depending on individual patient response, the dose may be increased to a maximum of 20mg once daily or decreased to a minimum of 5mg once daily. Adjustments are normally made early in treatment with assessment of response and tolerability at 2-4 weeks.
Special populations	The lowest effective dose of 5mg vortioxetine once daily should always be used as the starting dose in patients ≥65 years of age. Caution is advised when treating patients ≥65 years of age with doses higher than 10mg vortioxetine once daily for which data are limited. Dose adjustments may be considered in patients taking concomitant medications that induce or inhibit the P450 cytochrome system (see SmPC for details). The safety and efficacy of vortioxetine in children and adolescents aged less than 18 years has not been established.
Additional tests or investigations needed to identify suitable patients, or particular administration requirements	None.
Monitoring	Usual clinical practice for antidepressants (see Figure A5).

Other therapies likely to be administered at the same time as the intervention as part of a course of treatment	Antidepressant therapy may be provided in combination with a high-intensity psychological intervention, such as cognitive behavioral therapy or interpersonal therapy.

VAT: value-added tax; SmPC: Summary of Product Characteristics; MDE: major depressive episode; MDD: major depressive disorder

The ERG considers that the existing clinical pathways appear a reasonable interpretation of NICE CG90. However, as previously stated by the ERG, uncertainty exists surrounding the optimal position of vortioxetine within these pathways. Although the manufacturer has stated their preferred position of vortioxetine and defined their decision problem accordingly, the ERG considers that the focus on a restricted decision problem from the outset represents an important limitation of the submission. The ERG considers that the appropriate population and potential position of vortioxetine should have been more formally demonstrated by the manufacturer, based on a broader consideration of the evidence base for vortioxetine and other comparators. Consequently, by focusing entirely on the switch-population, the manufacturer subsequently excludes evidence from 22 of the 24 completed studies of vortioxetine, on the basis that these studies were not conducted in the population of interest. As a result, only the REVIVE and TAK318 trials are included on the basis that these provide direct evidence for the efficacy of vortioxetine in patients who switch from an initial SSRI or SNRI within an MDE and therefore address the decision problem. However, these 2 trials represent only 880 patients of the total of over 7,000 patients included within the completed set of vortioxetine studies.

3.3 Comparators

The NICE scope listed a broad set of comparators in line with the marketing authorisation in the UK for vortioxetine, including “SSRIs, tricyclic and tricyclic-related antidepressants, SNRIs, other antidepressant drugs and augmentation treatments”. The full list of comparators included in the NICE scope are summarised in Table 7.

Table 7 Comparators outlined in NICE scope

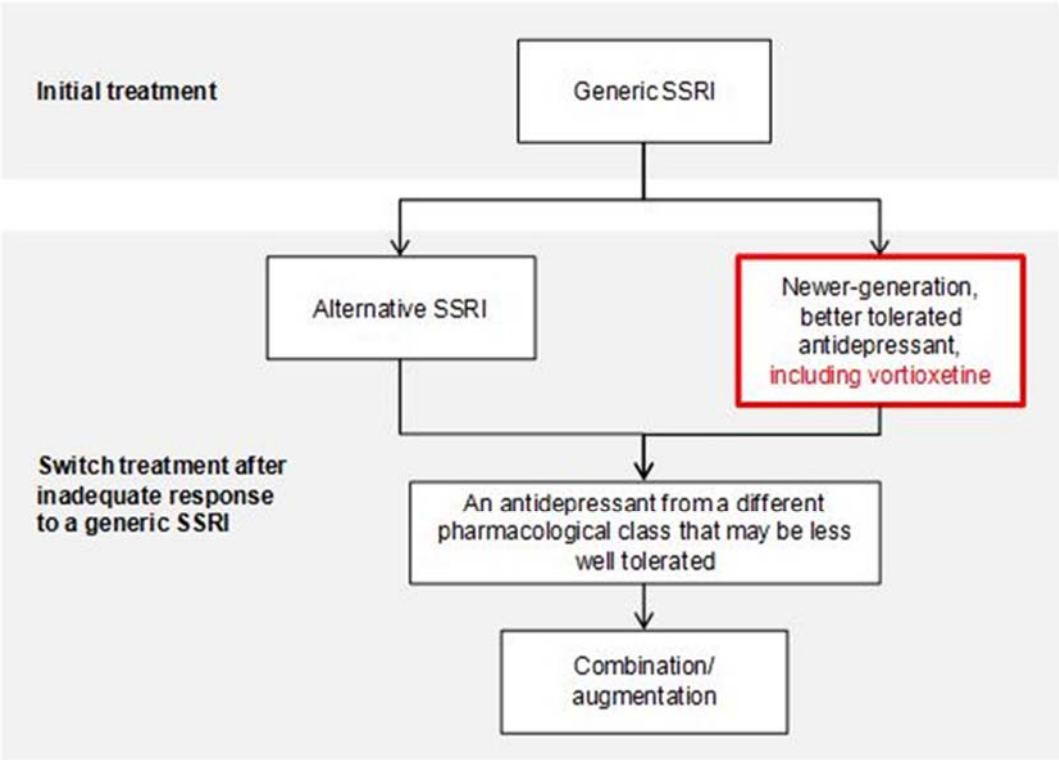
• Selective serotonin reuptake inhibitors (for example citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline)
• Tricyclic antidepressants (for example clomipramine, doxepin, imipramine, lofepramine, nortriptyline, trimipramine, and amitriptyline)
• Tricyclic-related antidepressants (for example mianserin and trazodone)
• Serotonin and noradrenaline reuptake inhibitors (for example venlafaxine, duloxetine and levomilnacipran)
• Other antidepressant drugs (for example agomelatine, mirtazapine, reboxetine and nonreversible mono-amine oxidase inhibitors [such as phenelzine])
• Augmentation treatments (for example, with an antipsychotic such as quetiapine)

The manufacturer significantly restricted the number of eligible comparators only including those which they considered represented alternatives in the proposed switch population (see Section 3.1 above).

The justification for the choice of comparators provided by the manufacturer was based on the recommendations of NICE clinical guidelines (CG90), clinical opinion and prescribing data in the UK(9), and was reported in MS section 2.7, pp.45-48. Clinical practice recommendations within CG90 state that when switching to another antidepressant, clinicians should consider switching initially to a different SSRI or a better tolerated newer-generation antidepressant and subsequently to an antidepressant of a different pharmacological class that may be less well tolerated, for example venlafaxine, an older TCA (e.g. amitriptyline) or an MAOI (e.g. phenelzine).

The manufacturer stated that the tolerability profile of vortioxetine, supported by the clinical efficacy data available within this population, was consistent with positioning within the category described in CG90 as: *"a better-tolerated newer-generation antidepressant"*. They further stated that proposed positioning of vortioxetine within existing pathways, summarised in Figure 4, also reflects the common practice of switching to an agent with a differing mechanism of action.

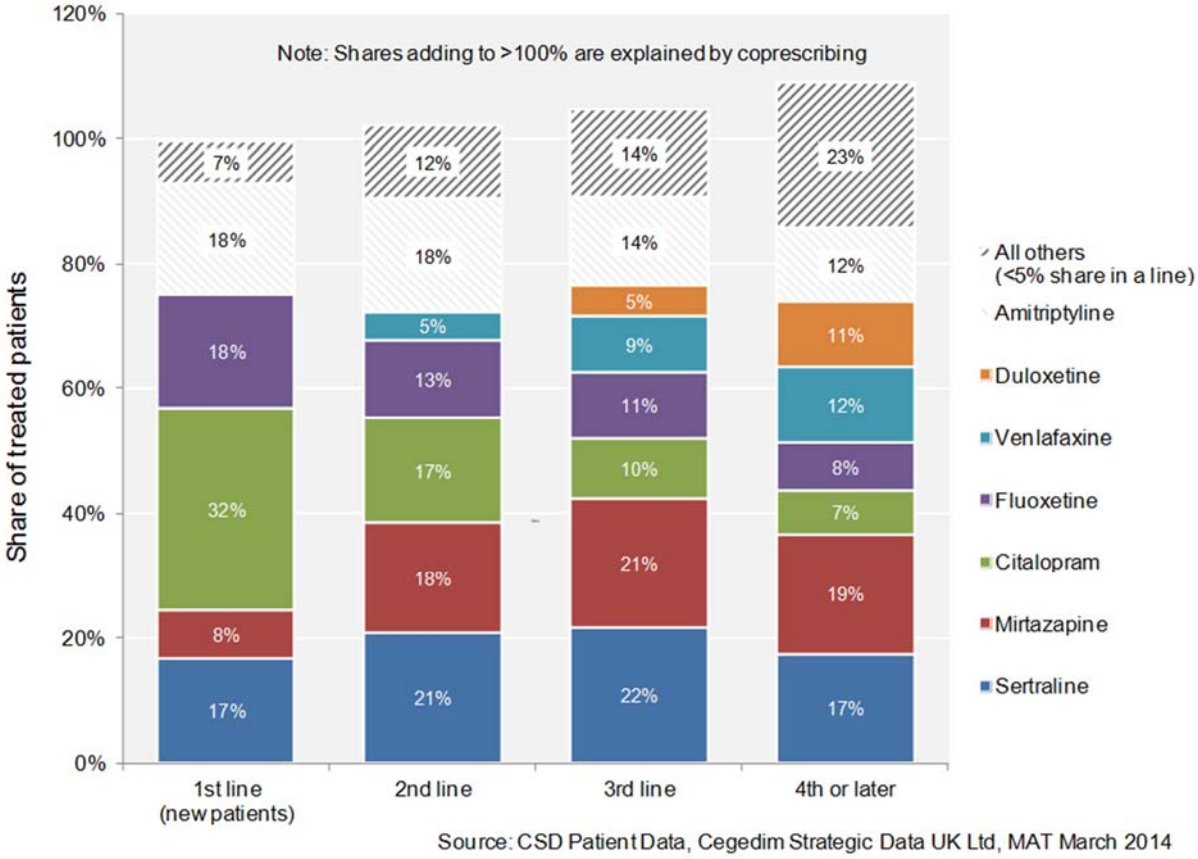
Figure 4 Modified treatment pathway from CG90 Guidance showing proposed positioning of vortioxetine



Based on these guidelines, the main comparators for vortioxetine as an initial switch therapy were stated by the manufacturer to be SSRIs and better tolerated, newer-generation antidepressants. The SNRIs (e.g. venlafaxine), tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) were argued by the manufacturer to be reserved for subsequent switches, as they *"may be less well tolerated"*. Combination/augmentation of antidepressants was not considered to be a relevant comparator in this submission by the manufacturer. These were excluded on the basis that they tend to increase the side-effect and drug interaction burden, and that this strategy should only normally be started in consultation with a consultant psychiatrist.

To further inform the selection of the most relevant comparators in the initial switch position, the manufacturer analysed market share data for the 12-month period April 2013-March 2014 to establish the market share for individual antidepressants by line of treatment. This is data is summarised in Figure 5.

Figure 5 Pharmacological treatments prescribed for patients diagnosed with depression, by line of therapy (from MS, Figure A10)



The manufacturer considered that the second-line (i.e. first-switch) data best reflected the initial switch population where the use of vortioxetine is proposed. Based on CG90, market share data and clinical opinion (i.e. the manufacturer concluded that the rates of prescribing for amitryptaline seemed implausibly high), the manufacturer specified the most appropriate comparators as (listed in descending order of second-line market share):

- Sertraline (SSRI)
- Mirtazapine (other)
- Citalopram (SSRI)
- Fluoxetine (SSRI)
- Venlafaxine (SNRI)

The manufacturer conducted systematic reviews (MS, Sections 6.1 and 6.7) to identify data to facilitate direct or indirect comparisons between vortioxetine and these comparators when used as second-line therapy. Studies were identified to support direct comparison with agomelatine (i.e. the comparator treatment in the REVIVE study) and indirect comparison with sertraline, citalopram and venlafaxine. No data were subsequently found to enable a valid comparison with mirtazapine or

fluoxetine in second-line use. While the manufacturer consider their inclusion would be desirable, they concluded that their absence was unlikely to be critical due to efficacy/cost considerations (i.e. they expected that fluoxetine would be dominated by other included SSRIs) and tolerability issues in a second-line setting (mirtazapine).

It is evident in subsequent sections of the ERG report that restricting the trial evidence to switching populations constrains both the evidence base considered as well as subsequent approaches to estimating comparative efficacy data to inform subsequent decisions. Similar concerns regarding the exclusion of potentially relevant evidence were reported by the Cochrane Depression Anxiety and Neurosis group as part of their comments on the draft scope. The group stated that:

> *"It's interesting to note that placebo is not mentioned as a comparator. On the one hand we would support this as a goal, as from a policy and clinical perspective, it is important to establish how vortioxetine compares with all other antidepressants. On the other hand, we would expect that excluding placebo controlled studies as comparators will lead to the exclusion of most randomized comparisons. Although we wouldn't challenge the key comparisons being made here, we do wonder if the scope should be widened to make best use of the available data on Vortioxetine, by considering placebo controlled trials as part of a network meta-analysis. This will enable NICE to consider all the comparative data to inform decision-making."* (Response to consultee and commentator comments on the draft scope, p.4)(10)

No change in scope was proposed in response to these comments on the basis that no treatment (placebo) is not established practice in the UK. However, importantly NICE also stated that:

> *"Placebo does not need to be included as a comparator in the scope in order for the company to be able to conduct a network meta-analysis including studies that compare the intervention with placebo"*. (Response to consultee and commentator comments on the draft scope, p.4)(10)

Importantly the existing clinical guideline (CG90) highlights additional uncertainties surrounding the interpretation of switching trials for comparator treatments as they often either include patients who may be expected to fare poorly on one of the treatments or employ a cross-over design. Furthermore, CG90 also 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. Consequently, in summarising the existing evidence and formulating guidance, the guideline group concluded that:

> *"Given the paucity of evidence from switching studies, evidence from primary efficacy studies in which antidepressants were directly compared were also considered. Caution is required in extrapolating from these studies to those whose illness has not responded to sequential trials of antidepressant drugs. Data from switching studies and head-to-head studies suggest that there may be a very small efficacy advantage for venlafaxine and escitalopram over other antidepressants. This advantage is too small to be clinically meaningful when all people with depression are considered together, but may be large enough to be clinically worthwhile in those who have not benefited from treatment with a first or second antidepressant. However, the current evidence is not sufficiently robust to form the basis of a recommendation".* (NICE CG90, p.479)*(1)*

The issues are further explored in subsequent sections of the ERG report.

3.4 Outcomes

The outcomes listed in the NICE scope were as follows:

- response to treatment (including response rate and time to response)
- relapse (including relapse rate and time from remission to relapse)
- severity of depression
- cognitive dysfunction
- remission of symptoms
- anxiety
- sleep quality
- hospitalisation
- mortality
- adverse effects of treatment (including adverse effects of treatment discontinuation)
- health-related quality of life

The manufacturer reported relevant data for most of these outcomes in the two trials of vortioxetine (REVIVE and TAK318). However, no data were reported on relapse, cognitive dysfunction, and sleep quality in REVIVE; in TAK318, no data on cognitive dysfunction, anxiety or sleep quality were presented, and no health-related quality of life outcomes, other than related to sexual dysfunction, were reported. The manufacturer stated that data for these outcomes had not been collected.

The primary outcome of the REVIVE trial was change from baseline in depression symptoms severity, which is a relevant outcome to the decision problem, although this was only measured in the short-term. The primary outcome of TAK318 was change from baseline in sexual dysfunction, which is of more limited relevance in the context of this appraisal.

The review of indirect evidence included significantly fewer outcomes, namely remission and withdrawal due to adverse events. The manufacturer stated that no other endpoints could be included in the review of indirect evidence as they were no reported consistently across all included studies. Although this justification appears appropriate, the absence of other outcomes (such as response rates) limits the relevance of the review of evidence comparing vortioxetine with several relevant comparators.

The MS cost-effectiveness model included some, but not all, of the effectiveness outcomes specified in the scope. In particular, the manufacturer did not include response data within the decision model.

3.5 Other relevant factors

The MS stated that no attempts were made to allow for equality considerations issues analytically, such as by applying alternative utility weightings. The MS noted that although people with intellectual disability are more likely to suffer from depressive episodes, this particular population may be at a disadvantage when it comes to receiving a specific diagnosis for depressive symptoms distinct from their other difficulties, and accessing the care they need (MS, Section 3, p.50). The ERG thinks that this approach was likely to be appropriate.

4 Clinical Effectiveness

The manufacturer submitted evidence on clinical effectiveness in the "switch" population based on four systematic reviews. This included three reviews of direct evidence for vortioxetine:

a) a review of RCTs of vortioxetine compared to active comparators in the switch population (MS, Sections 6.1 to 6.6) to evaluate efficacy;
b) a review of non-RCT evidence of vortioxetine in the switch population to evaluate efficacy (MS, Section 6.8);
c) a review of adverse events of vortioxetine (MS, Section 6.9).

Also submitted was a review of indirect comparative evidence, again in the switch population, including an indirect treatment comparison and network meta-analysis of antidepressant RCTs, to evaluate efficacy and safety (MS, Section 6.7).

Section 4.1 below summarises and critiques the methods of all four reviews. The first review identified two RCTs comparing vortioxetine to other active comparators, REVIVE and TAK318; these are critiqued in Section 4.2. The review of indirect treatment comparisons and its associated network meta-analysis are critiqued in Section 4.3. The review of non-RCT evidence did not identify any relevant studies, so this report does not consider this review in any detail (Section 4.4). The review of adverse events is critiqued in Section 4.5.

As discussed in Section 3 above, the ERG considers that the restriction to a switch population only is considerably narrower than the original scope specified and that assessment of the efficacy of vortioxetine should take account of the broader population by including evidence from studies of vortioxetine in non-switch populations. The ERG requested that the manufacturer provide data on all trials of vortioxetine compared to placebo and/or active comparators. The manufacturer supplied efficacy and adverse event data for their placebo controlled trials of vortioxetine. They also identified three published systematic reviews of vortioxetine, two submissions to regulatory authorities and one published indirect comparison of vortioxetine with other antidepressants, sponsored by the manufacturer. The ERG considers this evidence in the wider population to be of relevance when determining the clinical efficacy of vortioxetine; this evidence is critiqued in Section 4.6.

4.1 Critique of the methods of review(s)

4.1.1 Searches

The MS describes the search strategies used to identify direct evidence, indirect comparisons, non-RCT evidence and adverse events on the use of vortioxetine for the treatment of major depressive disorder. The strategies used for identifying the evidence are outlined in the main body of the submission with further details being provided in Section D of the Appendix.

Review of RCTs evidence

The manufacturer reported which bibliographic databases were searched. These include MEDILNE, EMBASE, PsycINFO and the Cochrane Library, among others. Trial registers including clinicaltrials.gov were searched, along with relevant conference websites.

The date when the searches were conducted, the time period covered by the searches and the total number of records identified were also provided. The search strategies used for each of the databases were reported. The search statements were combined appropriately and the correct Boolean notation was used. A PRISMA flowchart showed the total number of records identified and the contribution of each resource.

Reviews of adverse events and non-RCT evidence

The searches described for the review of RCT evidence would have identified RCT evidence, non-RCT evidence and adverse events studies, as no study type filter was applied to the searches. Consequently the comments on that search process apply to these sections too.

Indirect treatment comparison

The manufacturer reported which bibliographic databases were searched. A similar, but more limited, set of databases as in the review of RCT evidence was used. The full search strategies were provided in the appendix and it was noted that the strategy for EMBASE and MEDLINE was based upon a strategy used by a systematic literature review previously undertaken by the Agency for Healthcare Research and Quality (AHRQ), published in April 2012, assessing treatment for depression after unsatisfactory response to SSRIs.(11) The date the searches were conducted is given and it is reported that the search period covered was post 1980 onwards, although the reason for this restriction is not provided.

The ERG considers that the reporting of all search processes was clear, appropriate, and well documented.

4.1.2 Inclusion criteria

Review of RCTs evidence and review of non-RCT evidence

Eligibility criteria for the reviews of RCT and non-RCT evidence are reported in MS sections 6.2.1 and 6.8.1 respectively. Studies evaluating a licensed regimen of vortioxetine (5mg, 10mg, 15mg, or 20mg once daily) were eligible for inclusion. The population of interest was individuals with moderate-to-severe MDD who are experiencing an MDE, who have responded inadequately in terms of efficacy or tolerability to initial antidepressant treatment, and who require and want a switch to an alternative antidepressant. Eligible comparators were any antidepressants. There were no limitations in terms of eligible outcomes, and studies of any duration were eligible for inclusion.

Given the restricted population (see Section 3), the ERG considers the eligibility criteria for the reviews of RCT and non-RCT evidence to be appropriate.

Review of adverse events

Eligibility criteria for the review of adverse events are reported in MS section 6.9.1. Only studies that reported safety as the primary outcome were included. As with the reviews of RCT and non-RCT evidence, all licensed regimens were eligible, antidepressants were eligible as comparators, and the population of interest was switch patients with MDD.

These initial criteria did not lead to the identification of any studies of adverse events within a switch population. Therefore the selection criteria were broadened to include non-switch populations of adults with MDD (of any severity) experiencing an MDE. The manufacturer justified this decision based on the assumption that unlike efficacy, there is no reason to believe that the safety or tolerability of vortioxetine, or any other antidepressant, would differ by treatment line. Based on clinical advice, the ERG considers this assumption to be appropriate for safety, but not for tolerability, as patients intolerant of one medication may be more intolerant of others.

The study selection process and eligibility criteria appeared generally appropriate. However, the review of adverse events excluded studies that evaluated safety but did not report it as a primary outcome. Therefore there is a risk that relevant studies reporting adverse events may have been missed; as the manufacturer did not provide a record for the exclusion of studies, this risk is difficult to assess.

Indirect treatment comparison

The inclusion criteria for trials used in the indirect treatment comparisons were broadly similar to those for trials of vortioxetine. Included were adult patients with major depressive disorder, who for the current major depressive episode had demonstrated inadequate response to a previous treatment

(i.e. the "switch" population). Eighty percent or more of patients had to have received an SSRI or SNRI as first-line treatment. A range of antidepressant treatments (or placebo) were eligible, including all major, widely used treatments (citalopram, venlafaxine, sertraline, mirtazapine etc.). For the network meta-analysis inclusion criteria were more restricted, including only trials judged to be adequately randomised and blinded, based on quality assessment, which compared two or more antidepressants in the "switch" population.

The ERG considers that these inclusion criteria were appropriate.

4.1.3 Critique of data extraction

For all reviews data extracted included details of study design, participant characteristics, data relevant to risk-of bias, details of the treatments, outcomes (including changes in depression scores, response rates, remission rates, withdrawal rates and incidence of adverse events). The data extraction process appeared generally appropriate, although the manufacturer did not state whether attempts were made to minimise the risk of reviewer error and bias (for example, by independent checking or extracting data in duplicate).

4.1.4 Quality assessment

In all reviews the trials included were assessed for quality and risk of bias using seven questions from the NICE quality checklists. The assessment covered randomisation, allocation concealment, similarity of trial arms, blinding, imbalance across groups, outcome reporting and use of intention-to-treat analysis.

The ERG considers this to be an appropriate assessment of trial quality.

4.1.5 Evidence synthesis

Review of RCT evidence

No meta-analysis or evidence synthesis was performed for the trials of vortioxetine in the switch population as presented in the MS. The manufacturers justified this on the basis that only two trials (REVIVE and TAK318) were identified and they used different comparator treatments in different populations, so were unsuitable for synthesis. Instead, the results of the efficacy studies were summarised narratively and in tables. The ERG agrees that these two trials could not be reasonably synthesised and a narrative summary of the trials was appropriate.

Review of adverse events

For the review of safety studies, data from five of the six included studies were presented in aggregate to calculate the incidence of treatment emergent adverse events. No formal meta-analytic methods

were used to pool the safety studies. The ERG notes that this method is not ideal as it ignores differences in populations and characteristics across studies.

Indirect treatment comparison

Trials included in the indirect treatment comparison were pooled using network meta-analysis methods. The submission used two different approaches: a frequentist analysis using the Boucher method and a Bayesian analysis performed using WinBUGS. The ERG recognises that both approaches are standard methods, and have been used appropriately in the submitted analyses.

The ERG has concerns with some of the methods used in the network meta-analysis. The primary outcome was remission rate. No other efficacy outcomes, such as response rate or absolute changes in depression scores were included. On clarification, the manufacturer stated that no other efficacy endpoint was consistently reported across all studies and hence they could not be included in the network meta-analysis. Remission was defined as a HAM-D score of seven or less, or a MADRS score of ten or less. Because different trials reported different depression scales these scales may not be exactly comparable, and some patients may have achieved remission on one scale but not another. Also, one trial (Kasper) did not report remission rates, so this was calculated from HAM-D scores, assuming scores followed either a normal or gamma distribution. Sensitivity analyses were performed to investigate the effect of using both these distributions, which was appropriate. However, remission rates may have been inaccurately estimated if the data did not fit either of these distributions. On clarification, the manufacturer presented results of a network meta-analysis with standardised mean difference in depression scale as the outcome.

The main network analysis excluded trials with a placebo arm. The manufacturer justified this based on clinical advice that patients in placebo-controlled studies may be different from those in active-controlled studies. No further justification was provided. The ERG does not consider this to be a strong justification, because such differences would apply to all placebo-controlled trials, and including such trials would be unlikely to bias the network analysis results as a whole. The manufacturer did provide results of the network including placebo-controlled trials in Appendix 16.

The MS reported results as risk differences between treatments. The ERG does not consider the risk difference to be a suitable measure for comparing treatments as is it very sensitive to any heterogeneity in outcomes across trials. The manufacturers justified this choice because using the risk difference led to more conservative estimates for venlafaxine, bupropion and citalopram. The ERG agrees that results were more conservative in this case, but considers results based on odds ratios or relative risks to be more statistically robust. The manufacturers provided results based on odds ratios in an appendix, and results for relative risks in response to requests for clarification.

4.2 Critique of the trials of vortioxetine

The MS presented PRISMA flow diagrams for the review of effectiveness. Table 8 summarises the two RCTs identified by the review that compared vortioxetine to another antidepressant in the switch population.

Table 8 RCTs included in the submission

Study	Patient population	Regimen & duration	Comparator	Design	Follow-up duration	Primary outcome
REVIVE (14178A)(12)	Patients who have experienced an inadequate response to an SSRI or SNRI in their current MDE	Vortioxetine 10mg-20mg flexible dosing, 12 weeks	Agomelatine 25-50mg flexible dosing	Double-blind, international phase IIIb, parallel-group randomised trial	Efficacy: up to 12 weeks from baseline Safety: up to 16 weeks from baseline	Change from baseline in depression severity (MADRS total score) at week 8
TAK318(13)	Patients who are well-controlled on an SSRI but experienced treatment emergent sexual dysfunction	Vortioxetine 10mg-20mg flexible dosing, 8 weeks[a]	Escitalopram 10-20mg flexible dosing	Double-blind, multicentre phase IIIb, parallel-group randomised trial	Efficacy: up to 8 weeks from baseline Safety: up to 12 weeks from baseline	Change from baseline in sexual functioning (CSFQ-14 total score) after 8 weeks of treatment

[a] Participants who completed the 8-week treatment period entered a 1-week, double-blind taper-down period during which they received placebo. Escitalopram arm participants received 10 mg QD escitalopram during the taper-down period. Patients who prematurely discontinued during the double-blind treatment period were also offered taper-down study medication.
MADRS: Montgomery and Åsberg Depression Rating Scale; CSFQ-14: Changes in Sexual Functioning Questionnaire Short-Form

Study details and participant characteristics of vortioxetine trials were presented in section 6.3 of the MS (pp. 66-82), and efficacy results are reported in section 6.5 of the MS (pp. 87-99). Both included studies met the inclusion criteria. Reporting of study and participant characteristics appeared appropriate overall.

The ERG searched clinicaltrials.gov and the WHO ICTRP register and found no relevant ongoing vortioxetine trials in adult MDD switch populations.

4.2.1 Critique of the validity of the included trials

The ERG has several concerns with the validity of the included trials, particularly with how well they represent the UK population likely to receive vortioxetine. Both REVIVE and TAK318 had design

limitations that may have affected their external validity. The ERG has particular concerns regarding the choice of comparator, duration of follow-up and study power.

Comparators

Agomelatine is not currently recommended by NICE. The fact that agomelatine is the only comparator in REVIVE means that, in the context of this appraisal, the relevance of this trial is limited. The use of escitalopram is also limited in UK patients experiencing SSRI-induced sexual dysfunction, which limits the relevance of TAK318 to current UK practice.

Follow-up duration

The 12-weeks follow-up duration of REVIVE is short considering the duration of treatment that is recommended by NICE to achieve and consolidate remission. The manufacturer acknowledged this limitation (MS, p.136), but stated that this shortcoming was mitigated by the results of a relapse-prevention trial (study 11985A)(14) that found significantly lower relapse rates for vortioxetine compared with placebo (13% versus 26%, p=0.0013), over 24 weeks after 12 weeks' open-label treatment. The results of study 11985A should be interpreted with caution as it only included patients who were in remission at 10 and 12 weeks of open-label vortioxetine therapy, and therefore excluded responders-only, or patients who may have been in remission following a longer course of therapy. As with REVIVE, TAK318 had a short-term follow-up, therefore evidence of long-term efficacy of vortioxetine in switch populations is uncertain.

Study power

REVIVE was designed as a non-inferiority trial. Sample size calculations were based on a non-inferiority comparison of the treatment groups in the primary outcome, and 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. This means that REVIVE may not have been sufficiently powered to demonstrate that vortioxetine is superior to agomelatine. Therefore any inferences from this trial regarding the superiority of vortioxetine over agomelatine may not be reliable as they may be based on chance.

Additional concerns about the representativeness of the trial populations are reported in sections 4.2.3 (REVIVE) and 4.2.5 (TAK318) of this report.

4.2.2 Quality assessment and risk of bias

The manufacturer provided a quality assessment of the two included trials of vortioxetine (REVIVE and TAK318) in MS section 6.4, Tables B11 and B12 (MS, pp. 83-87). These are summarised in Table 9 below. The ERG mostly concurs with the manufacturer's assessment of quality for the two

trials, although there were minor concerns about baseline imbalances between treatment groups, withdrawal and missing data in both trials.

Prognostic factors were generally comparable across intervention and comparator groups in both trials, although small gender differences between intervention and comparator group were reported in REVIVE and TAK318. In REVIVE, the proportion of female participants was slightly higher in the vortioxetine arm (77.1%) compared with the agomelatine arm (72.3%). In TAK318, the proportion of female participants was slightly lower in the vortioxetine arm (56.9%) compared with the escitalopram group (60.8%). However, the ERG believes that these gender imbalances are relatively small and are unlikely to have introduced significant bias to the results.

In TAK318, ethnicity was also different between the 2 groups: 16.2% of participants reported being Hispanic or Latino in the escitalopram group, compared with 6.2% in the vortioxetine group. However, the ERG believes that this imbalance is unlikely to have a significant impact on the study results, as it is relatively small and involved few individuals.

Overall withdrawal rates were relatively high in REVIVE (23%) and TAK318 (22%). However there were no significant imbalances in withdrawal rates between treatment arms in either trials (21% in the vortioxetine group and 26% in the agomelatine group in REVIVE; 25% in the vortioxetine arm and 19% in the escitalopram arm in TAK318). Reasons for withdrawal were reported. Primary analyses of both trials used a mixed model for repeated measures (MMRM) to address the issue of missing data. Although this approach has limitations (the MMRM assumes that data are missing at random, and may therefore not fully compensate for informative but unobserved missingness), methods used to address missing data appear generally appropriate and the risk of attrition bias is likely to be low.

In REVIVE and TAK318, efficacy analyses were based on a modified intent-to-treat (mITT) analysis, using the full-analysis set (FAS) which comprised all randomised patients who took at least one dose of study medication, had a valid baseline assessment and at least one valid post-baseline assessment of the primary efficacy variable. This is not a strict intention-to-treat analysis according to the Cochrane Collaboration definition which requires the inclusion of all randomised patients. However, relatively few patients randomised in REVIVE (three in the vortioxetine arm and five in the agomelatine arm) and TAK318 (eight in the vortioxetine arm and 15 in the escitalopram arm) were not included in the FAS, therefore the risk of bias associated with the mITT analyses is likely to be low.

Table 9 Manufacturer assessments of quality for vortioxetine trials

	REVIVE	TAK318
1. Was randomisation carried out appropriately?	Yes	Yes
2. Was the concealment of treatment allocation adequate?	Yes	Yes
3. Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes
4. Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Yes	Yes
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	No
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
7. Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes

4.2.3 Characteristics of the REVIVE trial

REVIVE was a 12-week, international phase IIIb, non-inferiority, randomised, double-blind, parallel group, flexible-dose, active comparator study, that assessed the efficacy and safety of vortioxetine versus agomelatine in patients with MDD who had failed initial antidepressant therapy.

The patients were randomised equally (1:1) to flexible doses of either vortioxetine (10 to 20mg/day) or agomelatine (25 to 50mg/day). The starting doses were vortioxetine 10mg/day or agomelatine 25mg/day.

Participants were recruited from 71 psychiatric inpatient and outpatient settings in 14 countries (Austria, Belgium, Bulgaria, Czech Republic, Estonia, Germany, Italy, Lithuania, Poland, Romania, Russia, Spain, Sweden and the UK).

The study randomised 501 individuals to vortioxetine (255 participants) or agomelatine (246 participants). Included were adult patients (≥18 and ≤75 years) who had moderate to severe MDD (MADRS score ≥22) at screening and baseline and were candidates for a switch in the investigator's opinion. Patients needed to have responded inadequately to a maximum of one course of antidepressant SSRI or SNRI monotherapy that was prescribed to treat a single episode of MDD or recurrent MDD, according to DSM-IV-TR™ criteria.

A number of exclusion criteria were listed, including: current psychiatric disorder or Axis I disorder (DSM-IV-TR™ criteria), other than generalized anxiety disorder (GAD) and social anxiety disorder (SAD); MDD with post-partum onset or MDD with a seasonal pattern (DSM-IV-TR™ criteria); history of previous treatment resistant MDD; current diagnosis of alcohol or other substance abuse; significant risk of suicide, MADRS Item 10 (suicidal thoughts) score ≥5, or suicide attempt in

previous six months; currently receiving cognitive or behavioural therapy or systematic psychotherapy, or plans to start such therapy during the study. These criteria limit the extent to which the trial population may be representative of the UK switch population with moderate to severe MDD.

The primary outcome of REVIVE was change from baseline in MADRS total score at treatment week 8. Secondary efficacy outcomes included response (defined as ≥50% decrease from baseline in MADRS total score, or a CGI-I ≤2), and remission (defined as a MADRS total score ≤10, or a CGI-S ≤2).

MADRS is a ten-item diagnostic questionnaire used to measure the severity of depressive episodes. The overall score ranges from 0 to 60, with higher scores indicating more severe depression. The CGI-S scale is a seven point scale that rates the severity of a patients' mental illness. Scores range from one (normal) to seven (extremely ill). The CGI-I scale is a seven point scale assessing change relative to a baseline state in patient with mental disorders. Scores range from one (very much improved) to seven (very much worse), with scores of four indicating no change.

Other outcomes measures included safety endpoints (including adverse events and clinical safety laboratory tests), health-related quality of life and overall functioning. Follow-up was up to 12 weeks from baseline for efficacy outcomes, with an additional four weeks safety follow-up. All study outcomes assessed were reported in the submission and the trial CSR. A list of primary and secondary outcome measures is presented in Table B10 (MS, p.76).

Baseline characteristics of participants were reported in Table B8 (MS, p.75) of the submission, which is reproduced below (Table 10). The trial population appeared broadly representative of UK switch population. Baseline MADRS scores ranged from 22 to 43 points, which is consistent with patients experiencing moderate to severe MDD. However, all REVIVE participants (99.8%) were white Caucasian, which is unlikely to be reflective of the UK switch population. Approximately 76% of patients had received an SSRI, and 23% had received an SNRI as initial treatment. This is not representative of UK clinical practice, where first-line SNRI use is negligible, as acknowledged by the manufacturer. Participants from REVIVE were recruited almost entirely (97.2%) from outpatient psychiatric setting. The proportion of UK patients was small (approximately 7%). Variations in healthcare systems and management of MDD across different countries may limit the applicability of the trial results to the UK switch population with moderate to severe MDD.

Table 10 Study 14178A (REVIVE). Baseline characteristics of participants

Baseline characteristic	Vortioxetine	Agomelatine	Total
Number of patients: APTS[1], (FAS[2])	253 (252)	242 (241)	495 (493)
Mean (median) age (years)[1]	47.0 (48.0)	45.6 (46.0)	46.3 (47.0)
Sex (% female)[1]	77.1%	72.3%	74.7%
Race (% white)[1]	99.6%	100.0%	99.8%
Mean ± sd (median) duration of current episode (weeks)[1]	18.6 ± 10.4 (16.0)	19.2 ± 10.9 (16.0)	18.9 ± 10.6 (16.0)
Mean ± sd (median) number of previous episodes[1]	2.6 ± 2.1 (2.0)	2.4 ± 1.8 (2.0)	2.5 ± 2.0 (2.0)
Mean ± sd (median) MADRS total baseline score[2]	29.1 ± 4.4 (29.0)	28.7 ± 4.0 (28.0)	28.9 ± 4.2 (28.0)
Mean CGI-S ± sd (median) total baseline score[2]	4.4 ± 0.6 (4.0)	4.4 ± 0.6 (4.0)	4.4 ± 0.6 (4.0)
Mean ± sd (median) HAM-A total baseline score[2]	21.6 ± 6.3 (21.0)	21.4 ± 6.2 (21.0)	21.5 ± 6.2 (21.0)

[1] denotes analyses based on APTS (all patients treated set). [2] denotes analyses based on FAS (full analysis set). SD: standard deviation. See MS 6.3.6 for definitions of analysis sets

4.2.4 REVIVE trial results

Response and remission

REVIVE results for response and remission were reported in Table B14 (MS, pp.88-89) of the manufacturer submission, with further details presented in the study CSR. Response and remission rates, and corresponding adjusted odds ratios estimates from the manufacturer's logistic regression model are summarised in Table 11 below.

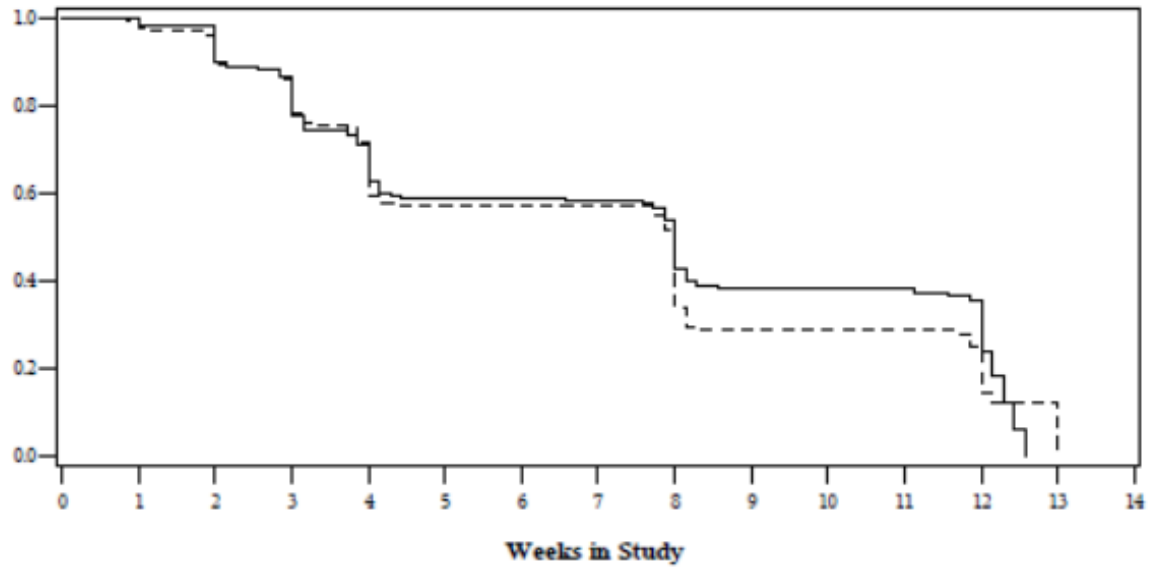
Table 11 REVIVE trial results: remission and response

Endpoint	Week 8				Adjusted OR (95% CI)	Week 12				Adjusted OR (95% CI)
	Vortioxetine		Agomelatine			Vortioxetine		Agomelatine		
	n	%	n	%		n	%	n	%	
Response										
MADRS	155	62%	114	47%	1.81 (1.26 to 2.60)	176	70%	135	56%	1.83 (1.26 to 2.65)
CGI-I	186	74%	140	58%	2.03 (1.38 to 2.96)	187	74%	154	64%	1.62 (1.10 to 2.39)
Remission										
MADRS	102	41%	71	30%	1.72 (1.17 to 2.52)	139	55%	95	39%	2.01 (1.39 to 2.90)
CGI-S	104	41%	78	32%	1.55 (1.07 to 2.25)	140	56%	106	44%	1.63 (1.14 to 2.33)
OR: Odds ratio; CI: Confidence interval										

Table 11 shows that the proportion of MADRS and CGI-I responders was statistically significantly higher in the vortioxetine arm compared with the agomelatine arm at weeks eight and 12. The proportion of MADRS and CGI-S remitters was also statistically significantly higher in the vortioxetine arm compared with the agomelatine arm at weeks eight and 12. Results were consistent across measures and follow-up points. The relatively wide confidence intervals reflect the uncertainty in true magnitude of the difference in efficacy between vortioxetine and agomelatine.

No data on time to response was provided in the submission. On clarification, the manufacturer conducted a post hoc time-to-event analysis of first response in REVIVE, along with a Kaplan-Meier plot of time to first response (defined as a ≥50% reduction in MADRS score), presented in Figure 6 below.

Figure 6 Kaplan Meir plot of time to first response (50% reduction in MADRS)



Treatment	N	No. of events	% events	Cox hazard ratio	P-value
Agomelatine	241	152	63.1	1.22	0.075
Vortioxetine	252	187	74.2	.	.
Total	493	339	68.8	.	.

Treatment —— PBO - - - - VOR

HTA UK ET_R50_178A Date and Time: 06MAR15 12:57

The results of this analysis showed a small separation between groups in favour of vortioxetine from 8 weeks, although the time-to-event analysis yielded a non-statistically significant result (p=0.075). This analysis should be interpreted with caution as it reports time to first response, and does not take into account patients who may have responded at early assessments but did not maintain response until the end of the study, as acknowledged by the manufacturer.

Severity of depression

Table 12 below presents differences in mean change from baseline between vortioxetine and agomelatine and shows that efficacy results were statistically significant across MADRS, CGI-S and CGI-I scores at treatment weeks eight and 12. REVIVE reported a reduction of 16.5 MADRS points (from 29.1 at baseline) in the vortioxetine group, and a reduction of 14.4 points (from 28.66 at baseline) in the agomelatine group, giving a mean difference of -2.16 (95% CI -3.51 to -0.81). The manufacturer stated that non-inferiority was established, as the upper bound of the 95% CI was below

the non-inferiority margin of +2 MADRS points versus agomelatine. The magnitude of the results was also similar between treatment weeks eight and 12 for each measurement scale. The manufacturer reported similar results from sensitivity analyses using LOCF and ANCOVA (Table B13, MS p88).

Table 12 REVIVE trial: Primary and secondary efficacy outcomes: scale score differences in mean change from baseline (MMRM)

Efficacy variables	Vortioxetine: difference to agomelatine	
	Week 8	Week 12
	Mean difference (95% CI)	Mean difference (95% CI)
MADRS total*	-2.16 (3.51 to -0.81)	-2.03 (-3.45 to -0.60)
CGI-S score	-0.30 (-0.48 to -0.11)	-0.27 (-0.47 to -0.07)
CGI-I score	-0.25 (-0.42 to -0.08)	-0.25 (-0.42 to -0.07)

*Primary outcome; MMRM: Mixed model for repeated measures

The ERG notes that MADRS results from MMRM analysis have relatively wide confidence intervals, therefore the magnitude of the mean difference estimate is uncertain.

Adverse effects of treatment

Table 13 presents a summary of adverse data.

Table 13 REVIVE trial: summary of adverse events reported over 12-week treatment period

Event	Vortioxetine (10mg-20mg)	Agomelatine (25mg-50mg)
	n=253	n=242
Adverse event		
Patients with TEAEs, n (%)	137 (54.2%)	127 (52.5%)
Patients with SAEs, n (%)	3 (1.2%)	4 (1.7%)
Patients with AEs leading to withdrawal, n (%)	15 (5.9%)	23 (9.5%)
Patients with TEAEs with an incidence of ≥5% in any treatment group (APTS):		
Nausea (%)	41 (16.2%)	22 (9.1%)
Headache, n (%)	26 (10.3%)	32 (13.2%)
Dizziness, n (%)	18 (7.1%)	28 (11.6%)
Somnolence, n (%)	10 (4.0%)	19 (7.9%)

TEAE: Treatment-emergent adverse event; SAE: serious adverse event; AE: adverse event.

Overall, most patients with TEAEs had TEAEs that were either mild or moderate. The incidence of patients with severe TEAEs was similar between the two treatment groups. The severe adverse events that were considered related to treatment and occurred in at least two patients in any treatment group

were (vortioxetine versus agomelatine): insomnia (4 patients versus 1 patient), headache (1 patients versus 3 patient), anxiety (1 patients versus 2 patient), tremor (2 patients versus 0 patient), and aggression (0 patients versus 2 patients).

The manufacturer stated that the overall incidence of sleep-related TEAEs was similar in the vortioxetine group compared to the agomelatine group (11.1% and 10.7%, respectively). Vortioxetine patients experienced insomnia (4.7%) more frequently than agomelatine patients (1.2%). Somnolence was less frequent in vortioxetine patients (4.0%) compared with agomelatine (7.9%).

The manufacturer did not present data on adverse effects of treatment discontinuation. In clarifications, they stated that the European Medicines Agency reported that there was no evidence of clinically relevant discontinuation symptoms that warranted a dose tapering of vortioxetine. The SPC stated that in short- and long-term placebo controlled studies, there was "no clinically relevant difference to placebo in the incidence or nature of the discontinuation symptoms after either short-term (6-12 weeks) or long-term (24-64 weeks) treatment with vortioxetine".

Other outcomes

No data on relapse rates, cognitive dysfunction or sleep quality (other than sleep-related adverse events) were presented. There were no deaths during the study and only one patient (in the agomelatine arm) was hospitalised.

Anxiety was measured using the Hamilton Anxiety Rating Scale (HAM-A). REVIVE reported a statistically significant difference in mean change from baseline in HAM-A total scores favouring vortioxetine compared with agomelatine at 8 weeks (MD -1.89; 95% CI -2.98 to -0.80) and 12 weeks (MD -1.93; 95% CI -3.04 to -0.81).

Several health-related quality of life measures were used, including EQ-5D, Sheehan Disability Scale (SDS), Depression and Family Functioning Scale (DFFS) and Work Limitations Questionnaire (WLQ). Statistically significant differences in favour of vortioxetine were reported for all outcomes at week 8 and all except the SDS family life domain and WLQ global score at week 12.

4.2.5 Characteristics of the TAK318 trial

Study TAK318 was a multicentre phase IIIb, randomised, double-blind, parallel group, flexible-dose, active comparator study that assessed the efficacy and safety of vortioxetine versus escitalopram in patients with well-treated MDD who were experiencing SSRI-induced sexual dysfunction.

Participants were recruited from psychiatry outpatient settings only, including 57 sites in the US and 9 sites in Canada. The study randomised 447 participants to 10-20mg/day of vortioxetine (225 participants) or 10-20 mg/day of escitalopram (222 participants).

Included were adults patients (≥18 and ≤55 years) who were currently being treated with SSRI monotherapy (citalopram, paroxetine or sertraline) for at least eight weeks for the treatment of an MDE according to the DSM-IV-TR™ criteria. Patients' depression was well treated and stable (CGI-S score ≤3); they were experiencing treatment emergent sexual dysfunction (CSFQ-14 total score ≤41 for women and ≤47 for men) considered to be attributable to the current SSRI monotherapy, and were suitable for a switch.

A number of exclusion criteria were presented, including: current psychiatric disorder; other comorbid conditions; current alcohol/substance abuse; current diagnosis or history of a psychotic disorder; significant risk of suicide, MADRS Item 10 (suicidal thoughts) score ≥5, or suicide attempt in previous six months; currently receiving cognitive or behavioural therapy or systematic psychotherapy, or plans to start such therapy during the study. The inclusion and exclusion criteria for this trial limit the extent to which the trial population may be representative of the UK switch population of patients with moderate to severe MDD with an inadequate response to initial antidepressant treatment. The primary outcome was change from baseline in sexual dysfunction (CSFQ-14 total score) after eight weeks of treatment. Follow-up was up to eight weeks from baseline for efficacy outcomes, with an additional four weeks safety follow-up. All study outcomes assessed were reported in the submission.

Relevant baseline characteristics of participants were reported in Table B9 (p.76) of MS, which is reproduced below (Table 14). On clarification, the manufacturer reported that 302 (71% of participants in the FAS) were in remission at baseline. Based on the information provided in the submission, it was unclear whether participants had received more than one course of antidepressant before entering the trial.

Table 14 Study TAK318. Baseline characteristics of participants

Participant characteristic	Vortioxetine (n=255)	Escitalopram (n=222)
Age, years Mean ± sd (range)	39.3 ± 10.0 (19–55)	40.2 ± 10.0 (19–55)
Sex		
Male	97 (43.1%)	87 (39.2%)
Female	128 (56.9%)	135 (60.8%)
Race		
Caucasian	178 (79.1%)	181 (81.5%)
Black	41 (18.2%)	35 (15.8%)
Asian	4 (1.8%)	3 (1.4%)
Other	2 (0.9%)	3 (1.4%)
BMI, kg/m^2, mean ± sd	27.5 ± 4.4	27.9 ± 4.4
CSFQ-14 total score Mean ± se (range)	36.1 ± 0.39 (21–47)	36.0 ± 0.40 (21–47)
MADRS total scoreMean ± sd (range)	7.9 ± 6.3 (0–34)	8.3 ± 6.5 (0–34)
CGI-S Mean ± sd (range)	2.0 ± 0.8 (1–3)	2.0 ± 0.8 (1–3)

sd: standard deviation; se: standard error; BMI: body mass index; CGI-S: Clinical Global Impression - Severity Scale; CSFQ-14: Changes in Sexual Functioning Questionnaire Short-Form; MADRS: Montgomery-Åsberg Depression Rating Scale.

4.2.6 TAK318 trial results

Response and remission

No results were reported for response and remission. TAK318 was conducted in switch patients whose MDD was well-treated and stable under previous SSRI treatment. MADRS and CGI-S baseline scores were therefore relatively low, and the lack of response and remission rates data from TAK318 is likely to be of limited relevance.

Relapse

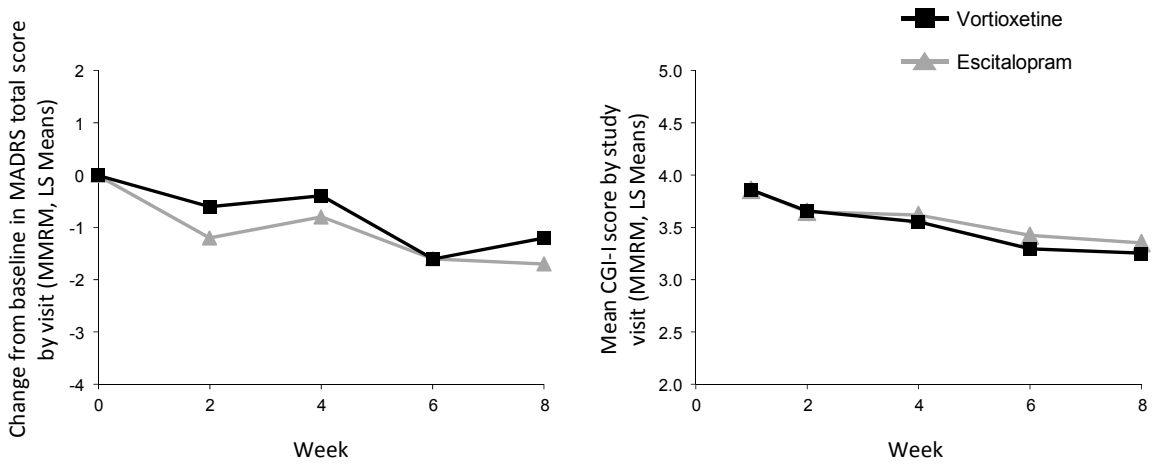
No relapse data were presented as part of the submission. A post-hoc analysis was conducted in response to request for clarification using MADRS score. Relapse was defined as MADRS total score of ≥22. [redacted]

Severity of depression

Efficacy results for depression symptoms were assessed using MADRS, CGI-S, CGI-I and POMS total scores. As can be seen from Figure B14 of the MS (reproduced in Figure 7 below), the mean MADRS and CGI-I total scores decreased slightly over time in both treatment groups, suggesting that

the improvement in depressive symptoms achieved with prior SSRI treatment at study entry was maintained or slightly improved overall. Results for CGI-S and POMS total scores were reported in the study CSR and showed a similar trend. Mean MADRS, CGI-S, CGI-I and POMS total scores were not significantly different between vortioxetine and escitalopram treatment groups at eight weeks.

Figure 7 Study TAK318. Changes in clinical measures of depressive symptoms over 8 weeks of treatment with vortioxetine and escitalopram



MMRM: Mixed model for repeated measures; LS: least-squares

Sexual dysfunction

Sexual functioning, measured as change from baseline in CSFQ-14 total score at treatment week eight was the primary outcome. At 8 week treatment the mean change from baseline was 8.8 (SE 0.64) in the vortioxetine arm and 6.6 (SE 0.64) in the escitalopram arm. There was a statistically significant difference in mean change from baseline in CSFQ-14 total score of 2.2 points after eight weeks of treatment which favoured vortioxetine (p=0.013).

The number of patients with a shift in CSFQ-14 from abnormal (defined as a CSFQ-14 total score ≤41 for women and ≤47 for men) at baseline, to normal (defined as >41 for women and >47 for men) was assessed as a secondary outcome. There was no statistically significant difference in the proportion of patients shifting from abnormal to normal CSFQ-14 score in the vortioxetine group (52.1%) compared with escitalopram (44.2%) at week eight (OR 1.37; 95% CI 0.93 to 2.03). The MS also reported a difference between arms on the CSFQ-14 subscales (MS, Figure B13, p98). There was statistically significant evidence in favour of vortioxetine on all subscales presented. Further secondary endpoint results and subgroup analyses were reported in the MS pp. 96-98.

Adverse effects of treatment

Of all participants who were randomized and received at least one dose of double-blind study medication, 283 patients (63.6%) experienced at least one TEAE during the Treatment Period (65.2% for vortioxetine and 62.0% for escitalopram). The overall incidence of TEAEs that caused discontinuation from study drug was 7.9% and lower in the escitalopram group (6.3%) compared with the vortioxetine group (9.4%).

The majority of TEAEs were mild or moderate in intensity. In the vortioxetine arm, the most common TEAE experienced by at least 5% of the participants were nausea (25.0%), headache (9.4%), dizziness (8.0%), and pruritus generalised (5.8%). In the escitalopram group, the most common TEAE were headache (7.7%), irritability (7.2), anxiety (5.4%), nausea (5.4%) and dizziness (5.0%).

Severe TEAEs were reported for 2.9% of participants overall and occurred with similar incidences in both treatment groups. SAEs were reported for three participants (1.3%) in the vortioxetine group and one individual (0.5%) in the escitalopram group. In clarifications, the manufacturer stated that no adverse event data associated with treatment discontinuation were available from TAK318.

Other outcomes

No deaths or hospitalisations were reported during the trial. No data on cognitive dysfunction, anxiety or sleep quality (other than sleep-related adverse events) were presented. No health-related quality of life outcomes, other than related to sexual dysfunction, were reported.

4.2.7 Conclusions from the review of vortioxetine RCTs

The review of efficacy studies identified two studies. Although both were conducted in switch patients, the populations of these trials differed significantly. REVIVE was conducted in MDD patients switching from initial SRI treatment (SSRI/SNRI) due to lack of efficacy, whereas TAK138 was conducted in patients whose MDD was well-controlled but were switched from an SSRI due to treatment-emergent sexual dysfunction. Therefore each trial only partly covered the switch population as defined by the manufacturer in the decision problem. Due to significant differences in populations, the ERG considers that the manufacturer's decision not to pool these two trials is appropriate.

The manufacturer concluded from the REVIVE trial that in MDD patients with an inadequate response to SSRI/SNRI treatment, switching to vortioxetine resulted in a significant and clinically relevant improvement versus agomelatine in change from baseline in MADRS total score at week eight. They stated that vortioxetine also showed a significant benefit versus agomelatine on the majority of secondary endpoints, and that vortioxetine and agomelatine were well tolerated with few treatment discontinuations.

Based on the TAK318 trial results, the manufacturer stated that vortioxetine is superior to escitalopram in improving sexual functioning in patients with SSRI-induced sexual dysfunction. They stated that both vortioxetine and escitalopram maintained and slightly improved the depressive symptoms seen with the prior SSRI treatment as assessed by MADRS, CGI-S, and CGI-I scores, and that vortioxetine was generally well-tolerated in this study.

REVIVE and TAK318 trials appeared well conducted. However, the designs of both studies raised a number of concerns. As mentioned above (Section 4.2.1), both trials included comparators of limited relevance to UK practice, and there were concerns about the representativeness of the trial participants to the UK switch population. Both studies had short-term follow-up, therefore the long-term efficacy of vortioxetine in switch patients is uncertain based on the evidence presented. In addition, the trials only evaluated the efficacy of vortioxetine 10-20mg/day, therefore the efficacy of the licensed 5mg/day regimen is uncertain. Finally, as REVIVE was designed as a non-inferiority trial, conclusions regarding the superiority of vortioxetine versus agomelatine may not be reliable.

4.3 Indirect treatment comparison

4.3.1 Critique of trials identified and included in the indirect comparison

The systematic review for the indirect comparison analysis identified 27 relevant studies; however 20 were excluded as not being in “switch” populations. The ERG accepts that these studies did not meet the inclusion criteria for the manufacturer’s analysis, but many of these studies may have matched the NICE scope population of adults with MDD.

The seven remaining studies were assessed for study quality. For five studies, all quality components were assessed as either adequate or unclear. One study (Rush 2006(5)) had one component (imbalances between groups) considered to be inadequate. The ERG generally agrees with the manufacturer on their quality assessment of these trials.

One study (Rosso 2012(15)) had two inadequate components (randomisation and blinding). Given the high potential for bias in this trial it was excluded from further analysis. The ERG accepts that assessment of bias was correct and so this exclusion was reasonable, but notes that, as Rosso 2012 was the only study to include duloxetine, this means no evidence comparing vortioxetine to duloxetine has been presented.

Table 15 Characteristics of studies included in indirect comparisons

Study	Interventions	Sample size	Mean age	% Female	Duration of current MDE (median months)	Number of previous MDEs (median)	Baseline HAM-D (mean)	Duration of previous treatment (mean, weeks)	% with prior SSRI/SNRI treatment	Primary outcome	Time of assessment (weeks)
REVIVE (16)	Vortioxetine 10-20mg Agomelatine 25-50mg	501	46	74.7	4.5	1.8	23.3*	24	100	MADRS	8
Kasper 2010(17)	Agomelatine 25-50mg Sertraline 50-100mg	177	44	73.5	3.1	2.0	26.5	NR	87.1	$HAM\text{-}D_{17}$	6
Olié and Kasper 2007(18)	Agomelatine 25-50mg Placebo	94	44.9	73	2.5	2.6	27.8	NR	81.9	$HAM\text{-}D_{17}$	6
STAR*D(5)	Bupropion 150-400mg Sertraline 50-200mg Venlafaxine XR 37.5-375mg	727	42	58.7	6.0	7.0	18.9	10	100	$HAM\text{-}D_{17}$	14
GSK 2009(19)	Bupropion 200-300mg Placebo	325	36	45.3	NR	NR	19	4	100	$HAM\text{-}D_{17}$	12
Lennox-Smith 2008(20)	Venlafaxine ER 75-300mg Citalopram 20-60mg	112	43	66.5	6.0	1.0	>31	8	100	$HAM\text{-}D_{21}$	12

** estimated by transforming MADRS; XR: extended release*

Baseline characteristics of the six included studies are presented in Table 15. Two trials (Kasper 2010, Olié and Kasper 2007) were not performed specifically in switch populations. Data on switch populations was extracted from a further post-hoc subgroup analysis of both these trials (Kasper 2013).(21) This subgroup analysis was also not strictly in switching patients as the subgroup considered was of previously treated patients, defined as patients who had been treated with antidepressants at least once in the year before randomisation. It was unclear whether all such patients were genuinely switching patients, or 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.

There is considerable diversity in baseline characteristics across trials. Of particular concern are the differences in baseline disease severity. STAR-D and GSK 2009 both have comparatively low baseline depression severity, below the standard threshold for moderate depression in the UK. Lennox-Smith, by contrast had very high baseline depression scores, as the population for this trial was restricted to patients with a HAM-D score over 31. Patients in STAR*D had generally much longer-term MDD, with more previous depressive episodes. As the effectiveness of antidepressants may vary by disease severity these differences could influence the outcomes of the indirect treatment comparison. Similarly the time of assessment varies from 6 to 14 weeks; again remission and withdrawal rates are likely to be time-dependent, so these differences could affect the results. The ERG therefore questions the validity of synthesising these heterogeneous trials as efficacy of treatments may not be consistent across diverse populations.

The primary network meta-analysis performed in the submission excluded the two placebo-controlled trials although these were included in a sensitivity analysis. Results on remission and withdrawal rates from the four remaining studies are presented in Table 16.

Table 16 Summary of outcomes from trials included in the main network meta-analysis

Study	Intervention	Number of patients	Remission			Withdrawal due to AEs		
			Number remitting	Rate (%)	Odds ratio and 95% CI	Number of withdrawals	Rate (%)	Odds ratio and 95% CI
REVIVE	Vortioxetine	252	102	40.5	1.63 (1.12 to 2.37)	15	5.9	0.6 (0.31 to 1.18)
	Agomelatine	241	71	29.5		23	9.5	
Kasper 2010	Agomelatine	80	NR	27.9 / 32.3 *	1.19 (0.61 to 2.24)	NR	2.6	0.28 (0.07 to 1.03)
	Sertraline	96	NR	24.5 / 27.3*		NR	11.3	
STAR*D	Sertraline	238	42	17.6	V vs S 1.54 (0.99 to 2.39) B vs S 1.27 (0.80 to 1.99)	50	21	V vs S 1.01 (0.65 to 1.56) B vs S 1.40 (0.92 to 2.14)
	Venlafaxine	250	62	24.8		53	21.2	
	Bupropion	239	51	21.3		65	27.2	
Lennox-Smith 2008	Venlafaxine	57 †	21	36.8	1.56 (0.99 to 2.45)	11	5.5	1.03 (0.68 to 1.57)
	Citolapram	55 †	15	27.3		11	5.3	

* Derived from normal/ gamma distribution assumptions; † Patients with MADRS >31

There are considerable differences in results across trials. Remission rates are comparatively low in STAR*D, and are much higher in REVIVE and Lennox-Smith. Similarly withdrawal rates are high in STAR*D and low in both REVIVE and Lennox-Smith. This further suggests that the trials and the populations they recruited may not be comparable.

Only the REVIVE trial found statistically significant evidence of a difference in remission rates between treatments, although comparisons between venlafaxine (XR) and sertraline, and between venlafaxine (XR) and citalopram were almost statistically significant. No trial found any statistically significant differences in withdrawal rates.

4.3.2 Critique of the network meta-analysis

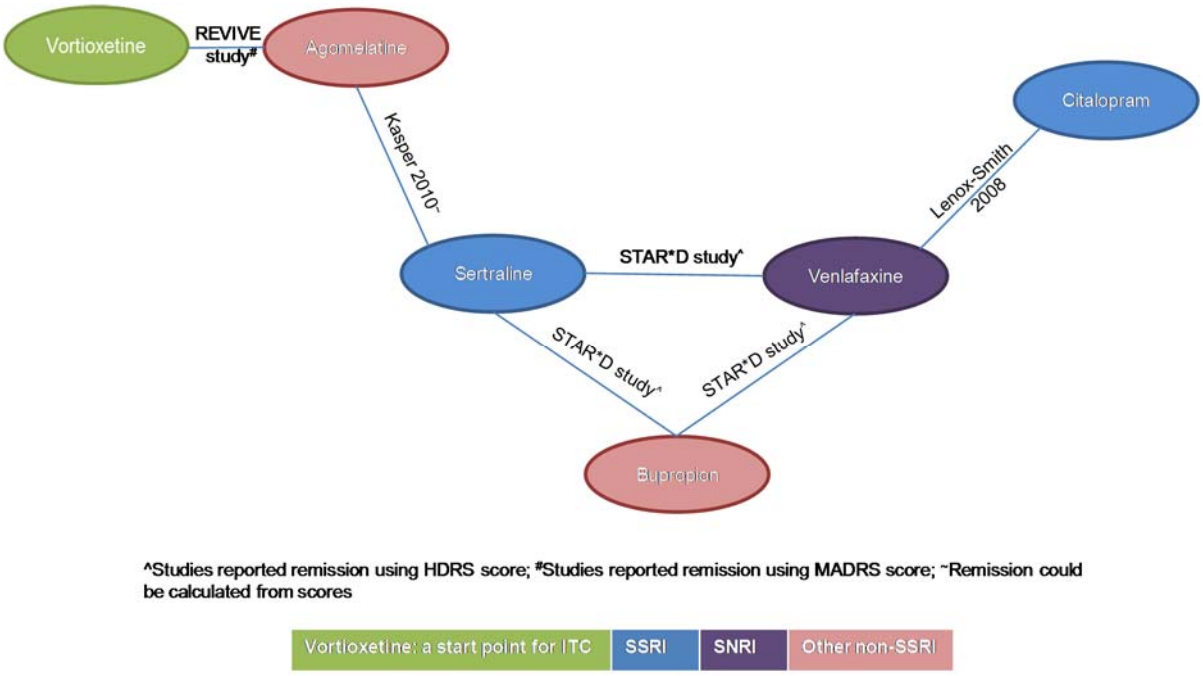
The network considered in the indirect treatment comparison is given in Figure 8 (taken from MS, p. 116). This network is based on only four trials, each arm of the network is informed by only one trial, and there are no “closed loops” in the network. The data for this network are therefore sparse, and comparisons between treatments are driven by the findings in each specific trial, which, as noted in Section 4.3.1, appear to be heterogeneous in their baseline characteristics and outcomes. The limited and heterogeneous nature of the data in the network means its findings may not be reliable.

Only one trial (REVIVE) included vortioxetine, so any comparison of vortioxetine with other treatments will be dependent of the results in that trial. Also only one trial included citalopram, so any

evidence to compare it with vortioxetine is weakened by limited evidence. Agomelatine may have lower efficacy and higher tolerability than other comparators (as suggested by Llorca et al.),(22) which may have biased the efficacy results of the indirect comparisons in favour of vortioxetine, and negatively biased the withdrawal results.

Removing any one trial from this network will leave an unconnected network for which no indirect comparison is possible. Given that the ERG questions whether the subgroup from the Kasper 2010 trial used in the network meets the inclusion criteria (because it does not appear to be specifically in a switch population) and that this trial is an essential link in the network, there is considerably uncertainty around the validity of any findings of this indirect treatment comparison. This should be considered when reading the remainder of this section.

Figure 8 Base-case network used in indirect treatment comparison



The indirect comparisons were performed for both remission rate and withdrawal rate as outcomes, and using both frequentist and Bayesian network meta-analyses. A summary of the results from the frequentist analysis is presented in Table 17. These results were based on assuming a normal distribution for the remission rate in the Kasper trial. Using a gamma distribution produced similar, but less conservative, results (see MS Appendix Table D49).

Table 17 Summary of the results of the frequentist network meta-analyses

Treatment	Remission rate			Withdrawal rate due to AEs		
	Rate (%)	Risk Difference vs vortioxetine (%)	95% CI	Rate (%)	Risk Difference vs vortioxetine (%)	95% CI
Vortioxetine	40.5	–	–	5.9	–	–
Agomelatine	29.5	-11	-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

The ERG considers that basing results on the risk difference is potentially inappropriate because it may be sensitive to heterogeneity across trials. In the appendix the submission also presented results based on the odds ratio using a Bayesian model. These are summarised in Table 18. Results were largely consistent with results based on risk difference.

Table 18 Summary of the results of the Bayesian network meta-analyses

Treatment	Remission rate			Withdrawal rate due to AEs		
	Rate (%)	Odds ratio vs vortioxetine (%)	95% CrI	Rate (%)	Odds ratio vs vortioxetine (%)	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

The manufacturer also provided results from a Bayesian network meta-analysis which included the two placebo-controlled trials excluded from the main analysis. As the results were broadly similar to those given above (see MS, appendix 10.17.1) this sensitivity analysis is not considered further. On clarification the manufacturer also provided results for analyses based on relative risks. Their findings were similar to the analyses based on odds ratios and risk differences, so they are not discussed further here.

On clarification the manufacturer provided results for analyses based on the standardised mean differences in depression rating scales. These results are summarised in Table 19. These results show efficacy in favour of vortioxetine, although the results are not statistically significant for comparisons with bupropion or citalopram. This finding appears to be dominated by the difference between vortioxetine and agomelatine from the REVIVE trial; other treatments seem broadly of similar efficacy.

Table 19 Summary of network meta-analysis reporting standardised mean differences

Treatment	**Standardised mean difference**	
	SMD vs vortioxetine	**95% CrI**
Agomelatine	0.304	0.109 to 0.449
Sertraline	0.542	0.186 to 0.898
Venlafaxine	0.559	0.161 to 0.957
Bupropion	0.609	0.210 to 1.008
Citalopram	0.630	0.186 to 1.074

For remission, frequentist and Bayesian analyses produced broadly similar estimates of remission rates. Neither analysis found any statistically significant evidence that vortioxetine was superior to any other treatment, other than agomelatine (from the REVIVE trial). For withdrawal due to adverse events both analyses found vortioxetine to have statistically significantly lower withdrawal rates than sertraline, venlafaxine (XR) and bupropion. The Bayesian analysis also found vortioxetine to be superior to citalopram. Results from risk difference and odds ratio analyses produced substantially different estimates of withdrawal rates, with the risk difference analysis producing lower rates. This difference suggests that estimates of withdrawal rate are highly sensitive to the method of analysis, and so are likely to be unreliable.

Each arm of the network included only one trial, and the network had no loops, therefore no assessment of heterogeneity or of network inconsistency was possible. The ERG notes that the substantial differences in baseline characteristics and outcomes across trials suggest that heterogeneity is likely to be present, reducing the reliability of the network analysis.

The manufacturer did explore consistency by comparing results from the network meta-analysis to the data from the trials themselves. Remission rates from the analysis and from the original trials are broadly similar, with rates being in general higher than reported in STAR*D, and lower than in Lenox-Smith. Withdrawal rates, however in the analysis were very different from those reported in

the trials. In particular the withdrawal rates on venlafaxine (XR) and citalopram were around 18% in the model (Table 17) but only around 5% in the Lenox-Smith trial (Table 16).

4.3.3 Conclusions of the indirect treatment comparison

The manufacturer concluded that the network meta-analysis showed that vortioxetine is more efficacious and better tolerated than a range of comparator antidepressants.

The ERG does not concur with this conclusion for a variety of reasons. Wide confidence intervals mean that there was little evidence of a statistically significant improvement in efficacy with vortioxetine compared to other treatments. The network meta-analysis used in the indirect treatment comparison was limited to only four trials. Removing any trial would leave the network unconnected. The ERG has concerns as to whether the Kasper trial should have been included as the subgroup analysis used from this trial was not specifically of a switch population.

There is also evidence of substantial heterogeneity in the populations included across the trial, for example in the severity of depression, and apparent heterogeneity in the results of each trial. The limited number of trials meant that formal assessments of heterogeneity and network inconsistency could not be performed. Given these concerns the ERG concludes that the indirect treatment comparison reported in the submission does not provide valid evidence form which any conclusion on the efficacy of vortioxetine may reasonably be drawn.

4.4 Non-RCT evidence

The manufacturers performed a search for evidence other than RCTs but did not identify any relevant studies. The ERG did not find any evidence that relevant non-RCT studies of efficacy in switch patients had been missed.

4.5 Adverse events

The manufacturer performed a search for safety trials of vortioxetine in the “switch” population. None were identified so the search was widened to include all safety trials of vortioxetine. This search identified six studies, of which five were one-year open-label extensions of short-term efficacy trials and were eligible for inclusion in a pooled analysis. This pool included a total of 2,587 patients who continued from short-term, randomised placebo controlled trials and received flexible doses of vortioxetine 2.5 to 20mg/day.

Given that the manufacturer had broadened the scope in the search for adverse event data to all vortioxetine trials, the ERG requested the manufacturer supply adverse event data from all placebo or active control trials they had performed. In response to this request the manufacturer provided data from 12 short-term placebo-controlled phase II/III studies, including one trial (12541A) that was

conducted in patients aged over 65 years. This short-term pool included a total of 3,904 patients treated with vortioxetine.

Table 20 provides a list of studies included in the two pools, and is presented below. None of the studies included were conducted in switch populations. Further details of the pools are presented below.

Table 20 Overview of pools for the evaluation of the safety and tolerability of vortioxetine

Pool	Number of studies included	Study ID
Short-term placebo controlled pool	12	11492A , 11984A, 305, 13267A, 315, 316, 14122A, CCT-002, 303, 304, and 317, and 12541A
Open-label long-term pool	5	11492C, 11984B, 301, 13267B, and 314

4.5.1 Short-term pool of placebo controlled studies

The results of the review of long-term continuation studies of adverse events were reported in a separate document(23) and summarised in clarifications. A summary of incidence of TEAE during the core treatment period (from first dose to last dose in the double-blind treatment period) for vortioxetine and placebo is presented in Table 21 below. The manufacturer stated that the 15mg dose will not available in the UK, but results for this regimen were presented here for the sake of completion.

The overall incidence of TEAEs was 64.2% in the vortioxetine group, and 57.9% in the placebo group. The Treatment-emergent adverse event (TEAE) with the highest incidence in all the active treatment groups was nausea. The incidence of nausea was 8.1% in the placebo group and 24.1% in the total vortioxetine group (relative risk 3.02; 95% CI 2.57 to 3.55). Headache, which was the TEAE with the highest incidence in the placebo group, had similar incidences in all the treatment groups (approximately 13%).

For the majority of the patients with TEAEs, the TEAEs were mild or moderate. The incidence of severe TEAEs was similar in the placebo group and in the vortioxetine total group (4.4% and 5.5%, respectively). In the vortioxetine dose groups, the incidence of severe TEAEs ranged from 4.1% in the 20mg group to 7.0% in the 5mg group.

The overall incidence of TEAE leading to withdrawal in the vortioxetine total group was higher in the vortioxetine group (6.0%) compared with placebo (4.0%), and higher in 15-20mg doses compared with 5-10mg doses (4.8% [5mg], 5.8% [10mg], 8.0% [15mg], and 7.0% [20mg]). The most common

TEAE leading to withdrawal during the core treatment period was nausea (vortioxetine: 1.1% [5mg], 1.4% [10mg], 3.3% [20mg]; placebo: 0.3%).

Table 21 Summary of TEAEs by dose (APTS)- Short-term pool

	Placebo (n = 1968)	**Vortioxetine 5mg (n = 1157)**	**Vortioxetine 10mg (n = 1042)**	**Vortioxetine 20mg (n = 812)**	**Total vortioxetine (n = 3460)**
Total adverse events	57.9%	64.7%	61%	65.1%	64.2%
Withdrawal due to adverse event(s)	4%	4.8%	5.8%	7.0%	6.0%
Severe adverse events	4.4%	7%	5.4%	4.1%	5.5%

Table 22 and Figure 9 from the manufacturer response to clarification show the estimated relative risk of adverse events (with 95% confidence intervals) for vortioxetine compared to placebo using the Cochran Mantel-Haenszel approach stratified by study. The analysis included data for vortioxetine 5 to 20mg/day from all studies in the MDD Short-term Pool. The analysis included adverse events with an incidence ≥2% for either vortioxetine 5 to 20mg/day or placebo.

Table 22 TEAEs with an incidence of 2% or more by preferred term, core treatment period (APTS) - Short-term pool

Type of treatment emergent adverse event	Percentage with event		Relative risk (95% CI)
	Placebo (n=1968)	Vortioxetine (n=3460)	
Event leading to withdrawal	3.6	5.5	1.54 [1.48 ; 1.59]
Severe AE	4.4	5.5	1.26 [1.22 ; 1.30]
Nausea	24.1	8.1	3.02 [2.57 ; 3.55]
Headache	13.2	12.9	1.06 [0.91 ; 1.22]
Diarrhoea	6.3	5.5	1.24 [0.99 ; 1.55]
Dry mouth	6.1	5.6	1.15 [0.92 ; 1.43]
Dizziness	5.8	5.3	1.14 [0.90 ; 1.44]
Nasopharyngitis	4.6	3.9	1.01 [0.78 ; 1.32]
Constipation	4	2.9	1.42 [1.04 ; 1.94]
Vomiting	3.9	1.1	3.55 [2.25 ; 5.60]
Somnolence	3.1	2.3	1.40 [0.98 ; 1.99]
Fatigue	2.9	2.7	1.11 [0.80 ; 1.55]
Insomnia	2.7	2.5	1.07 [0.76 ; 1.52]

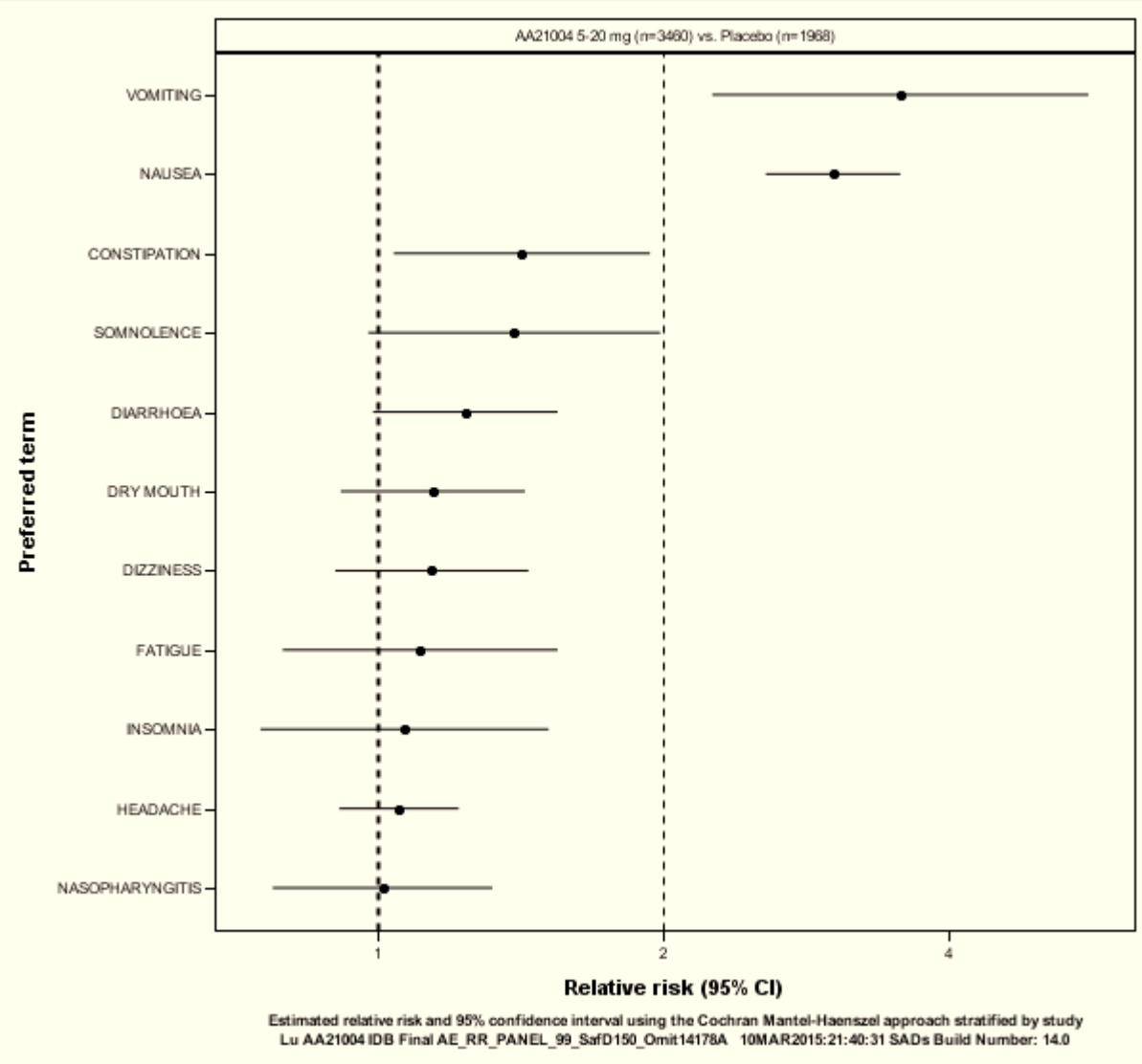


Figure 9 Forest plot of relative risk of TEAEs with an incidence of 2% or more by preferred term, core treatment period (APTS) - MDD Short-term pool

The analysis shows a statistically significant increase in risk of vomiting, nausea and constipation for patients taking vortioxetine. There was no other statistically significant difference in risk between vortioxetine 5 to 20mg/day and placebo in TEAE with an incidence of ≥2% in either group.

4.5.2 Open-label pool of continuation studies

The results of the review of long-term continuation studies of adverse events were reported in MS section 6.9.2 (pp.132-133) and in a separate document.(23) The review included six studies of vortioxetine in which safety outcomes were the primary endpoint. All were open-label extensions to short-term efficacy studies and were designed to provide 12-month safety data for regulatory

purposes. Five of those six studies were pooled in an analysis which provided input parameters for long-term side-effects in the economic model, were presented in the submission and in a separate document ("open-label long-term pool"). The sixth trial, which included 120 patients and was conducted in Japan, was excluded from the open-label long-term pool as it had not been completed at the time of the submission. The pool included a total of 2,587 patients who continued from short-term studies and received flexible doses of vortioxetine of 2.5 to 20mg/day. Of these patients, 1,391 (54%) were exposed to vortioxetine for ≥52 weeks. A small proportion of the pooled patients were in studies (301 and 11984B) that included a 2.5mg per day dose of vortioxetine. The manufacturer stated that it was not possible to split out the patients who received a 2.5mg dose because study 301 and 11984B were flexible dose studies in which patients received a daily dosage of between 2.5mg and 10mg, but as the proportion of patients receiving this dose was small it is not expected to affect the results.

A summary of incidence of TEAE during the core treatment period (from first dose to last dose of vortioxetine in the treatment period) is presented in Table 23 below. The overall incidence of TEAEs was 74.6%, and was higher in the 15-20mg dose group (78.9%) compared with the 2.5-10mg group (71.2%).

The overall incidence of TEAE leading to withdrawal was 8.1%, and the overall incidence of severe TEAE was 8.9%. The system organ classes (SOCs) with an incidence ≥20% were gastrointestinal disorders, infections and infestations, and nervous system disorders. The SOC with the highest incidence in both dose groups was gastrointestinal disorders (33% in the vortioxetine 2.5-10mg group and 42% in the vortioxetine 15-20mg group). The most common TEAE was nausea (20.4%).

Table 23 Incidence of TEAEs of ≥5% in the open-label, long-term pool, core treatment period

	Vortioxetine (2.5-10mg)		**Vortioxetine (15-20mg)**		**Vortioxetine (total)**	
	n	**%**	**n**	**%**	**n**	**%**
Number of patients (APTS)	1,443		1,144		2,587	
Patient-years exposure	1,097		775.4		1,873	
Patients with TEAEs	1,028	71.2	903	78.9	1,931	74.6
Patients with TEAEs leading to withdrawal	89	6.2	120	10.5	209	8.1
Patients with severe TEAEs	135	9.4	95	8.3	230	8.9

Core treatment period: from first dose to last dose of vortioxetine in the treatment period
TEAE: treatment-emergent adverse event.

4.5.2.1 Conclusions of the review of safety studies

No separate conclusions were provided for the review of safety studies, although the manufacturer stated that vortioxetine had a favourable safety profile in an overview of safety in relation to the decision problem (section 6.9.3).

The ERG agrees that based on the evidence presented, vortioxetine appears generally safe and tolerable in patients with MDD. The analysis of the pooled placebo controlled trials and the analysis of the pooled continuation studies included relatively large patient numbers and showed broadly comparable results. Although the incidence of adverse events was high in patients receiving vortioxetine, most AEs were mild to moderate in intensity and there was no conclusive evidence that these were dose-dependent.

However, both pooled analyses had limitations. All analyses from the pool of continuation studies were uncontrolled, and as such they are at high risk of confounding (notably due to the placebo effect). All continuation studies were one-year extensions, and nearly half of patients received vortioxetine for less than one year, which is significantly less than for patients for whom maintenance may be recommended for two years or more. Some relevant studies may have been missed, since only studies that reported safety as a primary outcome were included. Results from the pool of placebo-controlled studies showed that the rate of patients with adverse events was high for vortioxetine (64.2%) as well as for placebo (57.9%), indicating a high placebo effect, although there was a statistically significant higher risk in the incidence of some specific adverse events, particularly nausea and vomiting (approximately 3 and 3.5 times higher than placebo). These results are in line with those of the Pae 2014 review(24), which found that nausea and vomiting were some of the most common AEs reported, and had an incidence that was significantly higher in the vortioxetine than in the placebo group. They also found that nausea was the single most common AE reported as a reason for discontinuation of vortioxetine, and found that its frequency showed a trend toward a dose–response relationship. Compared with placebo, withdrawal due to adverse events was slightly higher for vortioxetine, particularly in higher treatment doses (15-20mg).

The MS and clarifications presented no evidence of adverse events when comparing vortioxetine with active comparators. The ERG concludes that the best data available are that from the two trials (REVIVE, TAK318) discussed above in Section 4.2. The indirect treatment comparisons provided by the manufacturer and in Llorca et al(22) both reported withdrawal due to adverse events, which can be considered a reasonable proxy for adverse events overall. These results are discussed in sections 4.3 and 4.6. Based on these results there is some evidence that vortioxetine has a better adverse event profile than other drugs, including venlafaxine and duloxetine. It may be less safe than agomelatine, although this conflicts with results from REVIVE.

4.6 Evidence syntheses of non-switch populations

As discussed in Section 3 the ERG questions the validity of restricting the analysis to "switch" populations. The evidence to suggest that there is a difference in treatment efficacy between initial use and switch use is limited,(1, 11) and restricted to showing that where patients are switching from an SSRI due to lack of response or intolerance, a non-SSRI may be more effective than another SSRI in these switching patients.(25) There is no evidence currently available to suggest that the relative efficacy of non-SSRIs differs between initial-use and switch populations. The ERG thinks that initial-use trials should be considered as providing relevant supporting evidence, particularly for the relative efficacy of non-SSRIs, given the limited nature of the evidence in the switch population identified by the manufacturer.

The ERG therefore requested that the manufacturer provide results from trials or meta-analyses of trials comparing vortioxetine to other active treatments and/or placebo in initial use and non-switch use populations.

In response the manufacturer:

1. Provided their own meta-analyses of trials comparing vortioxetine to placebo.
2. Reported the existence of two submissions to regulators:
   - Pharmaceutical Benefits Advisory Committee (PBAC, Australia)(26)
   - Canadian Agency for Drugs and Technologies in Health (CADTH)

   However, they did not provide these submissions; a redacted version of the PBAC report is in the public domain but the CADTH report has not been made public.

3. Identified four systematic reviews of vortioxetine:
   - Pae et al 2014(24)
   - Berhan and Barker 2014(27)
   - Citrome 2014(28)
   - Meeker et al 2015(29) (this review was published very recently and so was identified by the manufacturer subsequent to the MS and response to clarifications)
4. Identified one indirect treatment comparison sponsored by the manufacturer (Llorca 2015).(22)

The ERG thinks that all these analyses in non-switch populations provide relevant evidence on the efficacy of vortioxetine, and considers this evidence below.

4.6.1 Vortioxetine vs placebo

On request, the manufacturers provided results from meta-analyses of their short-term (6-8 weeks) placebo controlled trials of vortioxetine. The Pae et al.,(24) Berhan and Barker,(27) and Meeker et al(29) studies were systematic reviews designed to include all placebo controlled randomised trials of vortioxetine. The ERG considers that the search strategy, inclusion criteria and bias assessment processes in these reviews were generally appropriate. All papers reported the results of fixed or random-effects meta-analyses of the trials, for various outcomes. The Citrome paper was also a systematic review of placebo controlled randomised trials of vortioxetine, but did not include a conventional meta-analysis, so it is not considered further here.

There was considerable overlap in the trials included in the meta-analyses, they also reported slightly different outcomes. The manufacturer provided analyses by vortioxetine dose and a subgroup analysis using only non-US trials.

A summary of the results of the meta-analyses is given in Table 24. For data provided by the manufacturer and Meeker et al, results are presented for a 10mg dose of vortioxetine. The results are broadly consistent across analyses, as is to be expected given that they include a similar set of trials. All analyses show that vortioxetine is superior to placebo for all outcomes. Both the manufacturer's analysis and that of Berhan suggest a dose response relationship, with vortioxetine being more effective at higher doses, up to 20mg.

Table 24 Summary of meta-analyses of vortioxetine vs. placebo

Outcome	**Meta-analysis (effect estimate and 95%CI)**			
	Manufacturer*	**Pae et al.(24)**	**Berhan and Barker(27)**	**Meeker et al(29)**
Mean difference in change in MADRS	-3.53 -4.96 to -2.10	NR	-3.920 -5.258 to -2.581	-3.38 -4.89 to -1.87
Mean difference in change in CGI-I	-0.42 -0.59 to -0.26	NR	NR	NR
Standardised mean difference in depression score	NR	-0.217 -0.313 to -0.122	NR	NR
Response rate (odds/risk ratio)	1.84 1.44 to 2.35	1.652 1.321 to 2.067	NR	1.42 1.21 to 1.67
>50% reduction in MADRS	NR	NR	2.869 2.391 to 3.441	NR
Remission rate (odds/risk ratio)	1.59 1.23 to 2.04	1.399 1.104 to 1.773	NR	1.45 1.18 to 1.77
Discontinuation due to adverse event (odds ratio)	1.58 1.18 to 2.12[$]	1.530 1.144 to 2.047	NR	NR

*From clarification response, unless otherwise specified; [$] Llorca et al.(22)

4.6.2 Vortioxetine vs active comparators

The submission to PBAC, the Pae and Meeker meta-analyses all considered trials comparing vortioxetine to an active comparator. The trials included in these analyses were generally placebo controlled trials with an active reference arm, except for one randomised trial evaluating vortioxetine and venlafaxine in a head-to-head comparison which was included in the submission to PBAC.(30) In the original submission, the manufacturer criticised the use of these active references arms (MS, Sections 1.4.1 and 2.8.2). Active reference arms are included in trials of antidepressants to ensure that patients are responding to therapy. An active reference should be a drug of proven superiority over placebo, so it can be used to check whether the trial has successfully treated patients by confirming a difference between the active reference and placebo. Patients known to be non-responders to the active reference are excluded from this arm, so the active reference arm may include patients more likely to respond to treatment. The manufacturer therefore claimed that any comparison of vortioxetine with an active reference may be biased towards the active reference.

The ERG recognises the potential for bias because the comparison between vortioxetine and active reference is not truly randomised and there is the potential for patients to differ in likely response to treatment between arms. However, the ERG does not consider this risk of bias to be sufficiently substantial to reject such comparisons altogether, particularly as there is a reasonably large number of such trials. The results of these comparisons discussed in the Pae and Meeker analyses, and the PBAC submission are therefore considered below.

4.6.2.1 Pae et al.

In Pae et al. meta-analyses comparing vortioxetine to active comparator for standardised mean difference, response and remission rates were presented. The analysis did not distinguish between the different comparator treatments. Based on the references in the paper the ERG has reanalysed the data, grouping trials by the comparator included in each trial.

Figure 10 shows the results based on data from the trials in the Pae meta-analysis for standardised mean difference in depression score between arms. Figure 11 shows the results for response rate, and Figure 12 for remission rate. In all three cases vortioxetine is superior to agomelatine based on the trial of Dragheim.(31) This trial appears to be an early presentation of the results of the REVIVE trial, but this could not be confirmed because the trial was reported only as a conference presentation which the ERG could not access. There is no evidence of a difference between vortioxetine and venlafaxine, based on one trial. Vortioxetine is consistently and statistically significantly inferior to duloxetine for all three outcomes. The potential that results were biased in favour of duloxetine given the potential for bias discussed above must be considered, however there is no evidence of such a bias in favour of venlafaxine.

Figure 10 Forest plot of standardised mean difference based on data from Pae et al.

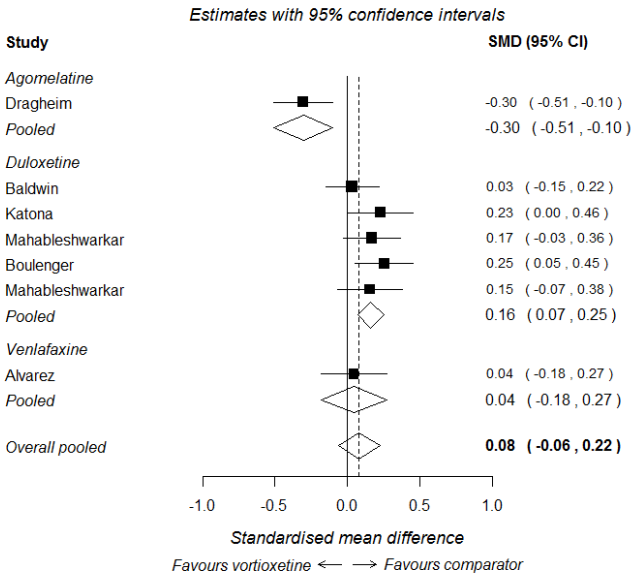


Figure 11 Forest plot of response rate based on data from Pae et al.

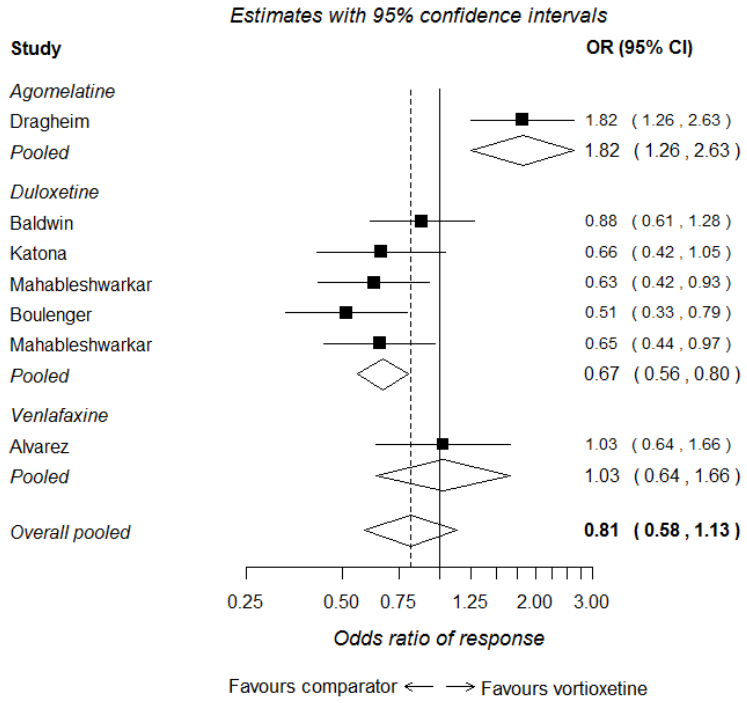
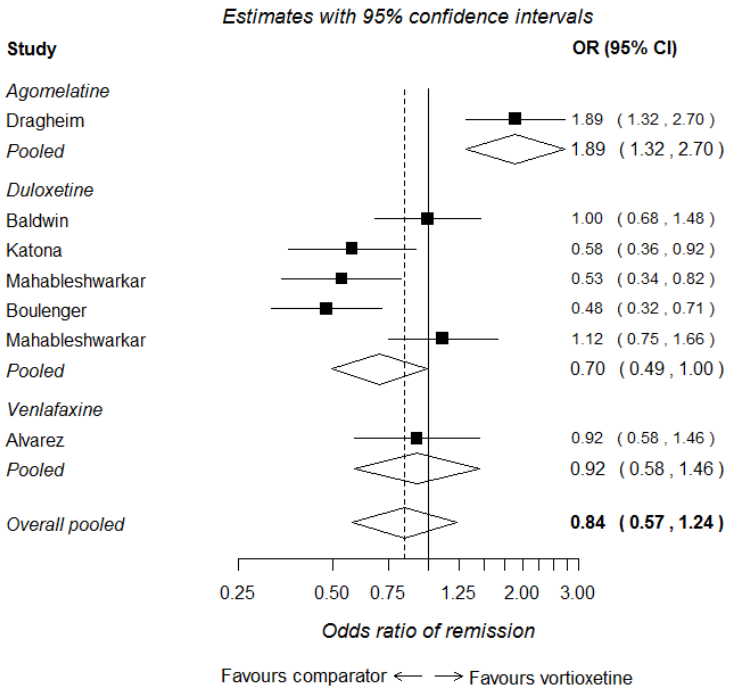


Figure 12 Forest plot of remission rate based on data from Pae et al.



4.6.2.2 Meeker et al.

Meeker et al compared vortioxetine to SNRIs, using a similar set of trials to Pae et al. The paper did not distinguish between different SNRIs and reported results separately for different doses of vortioxetine. Given this, it was not possible to directly compare results from this paper to other analyses, but generally SNRIs were found to have a better response rate than vortioxetine and had greater reductions in MADRS scores. Data on remission were too sparse to draw any conclusions.

At higher doses (15mg and 20mg) vortioxetine had lower rates of withdrawal due to adverse events than SNRIs, but there was no difference in the incidence of serious adverse events.

4.6.2.3 PBAC submission

The submission to PBAC analysed vortioxetine trials with either venlafaxine of duloxetine as active comparators. Two venlafaxine trials were included. Five duloxetine trials were included, but data were redacted for all but one. There was overlap between the trials included in the PBAC submission and Pae et al. Table 25 presents a summary of the results included in the submission.

Table 25 Summary of results in the PBAC report

Outcome	Active comparator (Mean difference and 95% CI)*	
	Venlafaxine (Two trials)(30, 32)	**Duloxetine (one trial)(33)**
Mean difference in MADRS	-0.44 -2.20 to 1.32	2.50 0.41 to 4.59
Mean difference in HAM-D17	-0.35 -2.07 to 1.37	NR
Mean difference in HAM-D24	-0.22 -2.35 to 1.91	2.10 0.04 to 4.16
*Positive values indicate a difference favouring the active comparator		

As for the Pae et al analysis, there was no evidence of any difference in efficacy between vortioxetine and venlafaxine, but duloxetine was superior to vortioxetine. The PBAC submission reported that treatment withdrawals due to adverse events were significantly more common in the venlafaxine arm than the vortioxetine arms, and found no statistically significant differences between vortioxetine and duloxetine. The PBAC concluded: “the claim of non-inferiority of vortioxetine compared to duloxetine was not adequately supported” (PBAC, p.12),(26) suggesting that the results of redacted duloxetine trials were consistent with that for which data were reported.

The results of the Pae meta-analysis and the PBAC submission appear to be consistent.

4.6.3 Indirect treatment comparison

On request for clarification the manufacturer identified a further indirect treatment comparison (Llorca et al).(22) This was a systematic review and network meta-analysis, sponsored by the manufacturer. The review sought to identify all placebo controlled trials of the following drugs: vortioxetine, agomelatine, desvenlafaxine, duloxetine, escitalopram, sertraline, venlafaxine, vilazodone. The outcomes considered were efficacy (in terms of standardised mean difference in depression scales, remission and response rates) and tolerability (in terms of odds ratio for withdrawal due to adverse effects). Trials were pooled using random-effects meta-analyses for each drug. Indirect treatment comparisons (via placebo) were performed using linear meta-regression models. The ERG considers that the review process and statistical methods used in this review appeared generally appropriate, but notes that no quality assessment of the included trials was reported.

The results of the indirect treatment comparison performed by Llorca et al are given in Table 26.

Table 26 Indirect treatment comparison by Llorca et al.

Outcome	Results vs vortioxetine (Standard error)						
	Agomelatine	**Desvenlafaxine**	**Duloxetine**	**Escitalopram**	**Sertraline**	**Venlafaxine**	**Vilazodone**
Efficacy at 2 months (SMD: <0 favours vortioxetine)	-0.156 (0.113)	0.025 (0.803)	0.090 (0.419)	-0.054 (0.695)	-0.037 (0.832)	0.124 (0.328)	-0.245 (0.111)
Withdrawal (Odds ratio: <1 favours vortioxetine)	1.769 (0.030) *	0.578 (0.035) *	0.752 (0.262)	0.671 (0.275)	0.299 (0.008) **	0.469 (0.009) **	0.640 (0.181)
Response rate (Odds ratio: >1 favours vortioxetine)	1.045 (0.815)	1.153 (0.364)	0.893 (0.514)	0.843 (0.523)	0.772 (0.575)	0.789 (0.353)	0.975 (0.934)
Remission rate (Odds ratio: >1 favours vortioxetine)	1.220 (0.470) **	1.029 (0.852)	0.894 (0.526)	0.990 (0.981)	NA	0.689 (0.444)	0.983 (0.952)

** p-value 0.01 – 0.05; ** p-value <0.01*

This analysis generally found no statistically significant evidence of a difference in efficacy between vortioxetine and any other treatment. The exception was that vortioxetine had a higher remission rate than agomelatine. Vortioxetine generally had a lower withdrawal rate than other treatments, although the results were only statistically significant for desvenlafaxine, sertraline and venlafaxine. Agomelatine had a lower withdrawal rate than vortioxetine.

This analysis was not performed in a “switch” population, however the findings of this analysis are consistent with those from the indirect treatment comparison in the switch population submitted by the manufacturer (except for agomelatine and withdrawal).

4.7 Conclusions of the clinical effectiveness section

The manufacturer’s submission focused on trials in patients who had previously received antidepressants for a current MDE but were switching treatment to vortioxetine due to lack of response or adverse events; the “switch” population. As discussed in Section 3The ERG notes that this is a substantial restriction when compared to the original scope, which specified a general population of all adults with MDD.

The restriction to a switch population meant that only two trials comparing the efficacy vortioxetine with other antidepressants were submitted. Both trials were generally well conducted but there were a number of issues related to their design and their population which may limit the applicability of their

results to UK practice. The REVIVE trial comparing vortioxetine to agomelatine found vortioxetine had greater reduction in depression scores and lower withdrawal rates due to adverse events, but this result may not be reliable as this was a non-inferiority trial and was not powered to detect superiority of vortioxetine. It should also be noted that agomelatine is not approved for use in the UK as no evidence for its efficacy has been submitted.

The TAK318 trial included patients who were switching due to sexual dysfunction but had responded to initial treatment. The ERG notes that this is a very narrow and specific population, so the TAK318 trial provided little information on the broader population who might take vortioxetine. It did find that vortioxetine reduced sexual dysfunction symptoms when compared to escitalopram. However it is not clear whether this finding is specific to vortioxetine, or whether any non-SSRI would have a similar beneficial effect.

The restriction to switch populations meant that the indirect treatment comparison in the submission included only four trials in its main analysis. The ERG has considerable concerns over the validity of the network analysis because of the high apparent diversity in the populations across trials, including very different patients and severities of depression. For one trial (Kasper) the analysis used a subset of patients who had been treated in the past year. This is not the same as patients who were switching treatment, so the ERG does not think this trial should have been included; without it, there is no connected network, rendering the analysis invalid. As such the ERG does not think the results of the indirect treatment comparison are valid, but notes that the analysis found no convincing evidence of difference between vortioxetine and other treatments in terms of remission rate (except for agomelatine). There was some evidence that vortioxetine may have lower withdrawal rates, but the high apparent heterogeneity across trials means the validity of this finding is questionable.

The manufacturers did not identify any trials reporting safety data in the switch population other that data reported in REVIVE and TAK318. The manufacturers therefore expanded the review of safety data to non-switch populations. This review identified 12 short-term placebo controlled trials and five longer-term open label trials of vortioxetine. In the placebo controlled trials severe adverse event rates and adverse events leading to withdrawal were more common on vortioxetine than placebo. However, based on the evidence presented, vortioxetine appears generally safe and tolerable in patients with MDD. Both the analysis of placebo controlled trials and the analysis of continuation studies included relatively large patient numbers and showed broadly comparable results. Although the incidence of adverse events was high in patients receiving vortioxetine, most AEs were mild to moderate in intensity and there was no conclusive evidence that these were dose-dependent. The manufacturer did not present any safety comparisons of vortioxetine with any active comparators in the broad MDD population, so the safety profile of vortioxetine compared to other antidepressants is uncertain.

Given the limited nature of the data in the switch population the ERG considers that data in non-switching and initial-use populations should be considered. Although such data is not in the switch population it is relevant to the broader population of all patients with MDD specified in the NICE scope. The manufacturer justified excluding trials of non-switching populations by claiming that treatment efficacy in a switch population may be different from in initial use. The ERG considers that the evidence submitted to justify this claim was limited, as it refers only to patients who had previously used an SSRI, where switching to another SSRI may be less effective than a non-SSRI treatment. No evidence was presented to suggest that the relative efficacy of non-SSRIs may vary between initial and switch use, and no evidence was specific to vortioxetine. The ERG therefore concludes that this restriction was inappropriate and evidence on non-switch populations is relevant when examining the efficacy and safety of vortioxetine.

Direct evidence comparing vortioxetine to other active treatments in non-switching populations was primarily available from placebo controlled trials with active reference arms. The manufacturer criticised the use of active references because they are not true randomised comparisons and patients known to be non-responsive to the reference are excluded, possibly biasing results in favour of the active reference. While the ERG accepts the potential for such bias it does not consider this potential bias to be substantial enough to exclude these trials.

Data on trials with active reference arms were available from the meta-analysis of Pae et al, and the manufacturer's submission to PBAC. Both found no evidence of any difference in efficacy between vortioxetine and venlafaxine, based on two trials. There was evidence that vortioxetine was inferior to duloxetine in terms of reducing depression scores, response and remission. While there is a possibility of bias in favour of duloxetine in these analyses it is not clear whether any bias would be sufficient to completely explain this inferiority.

An indirect treatment comparison in non-switch populations was performed by Llorca et al, based on placebo controlled trials of a number of antidepressants. This analysis found no evidence of any difference in efficacy between vortioxetine and other treatments, but there was some evidence to suggest that vortioxetine had a lower withdrawal rate due to adverse events than some treatments, including sertraline and venlafaxine. While this is an indirect analysis, and not conducted in a switch population, the number of trials included in this analysis suggests that this may represent the most reliable evidence for comparing vortioxetine to other treatments.

In summary, the manufacturer's restriction to trials in a "switch" population meant that very limited evidence was presented in the submission and the ERG considers the indirect treatment comparison, in particular, to be unreliable. Trials in the more general, non-switching, population provide more

data, but with the possibility that treatment effects may differ from those patients switching treatments, particularly when comparing vortioxetine to SSRIs. Direct comparisons of vortioxetine to other treatments are limited because they are placebo-controlled trials with active reference arms and so there is potential for bias due to them not being truly randomised. These trials, however, suggested that vortioxetine may be inferior to duloxetine, and this possibility cannot be dismissed entirely even with the potential for bias. Indirect comparisons of treatments suggested that vortioxetine had similar efficacy to other drugs, but with a possibility lower withdrawal rate due to adverse events.

The ERG concludes, based on the totality of the evidence, that vortioxetine is likely to be of similar efficacy to other antidepressants, but may be superior to agomelatine and inferior to duloxetine. Vortioxetine appears to have a lower withdrawal rate due to adverse events than most other treatments, and so may have a better overall safety profile, however data on adverse events with vortioxetine, particularly when compared to other antidepressants, are too limited to draw any firm conclusions on the safety of vortioxetine.

5 Cost Effectiveness

This section focuses on the economic evidence submitted by the manufacturer and the additional information provided in response to the ERG points for clarification. The submission was subject to a critical review on the basis of the manufacturer's report and by direct examination of the electronic version of the economic model. The critical appraisal was conducted to highlight key assumptions and areas of uncertainty. Section 6 presents additional work undertaken by the ERG to further explore these aspects.

The manufacturer's economic submission included:

- A description of the systematic literature review conducted to identify existing published evidence on the cost-effectiveness of vortioxetine relevant to the stated decision problem concerning the management of adults experiencing an inadequate response to a prior SSRI or SNRI for an MDE (MS, Section 7.1).

- A report on the de novo analysis undertaken by the manufacturer. The report described the patient population, the model structure and the associated treatment pathways assumed before and after a possible switch to third and later lines of treatment (MS, Section 7.2); the clinical parameters used in the economic model (MS, Section 7.3); the measurement and valuation of health effects and the quality-of-life data used in the cost-effectiveness analysis (MS, Section 7.4); the resource use identification and the parameters used in the model (MS, Section 7.5); the sensitivity analyses undertaken (MS, Section 7.6); and the cost-effectiveness results for the base-case and sensitivity analyses (MS, Section 7.7).

- An electronic copy of the manufacturer's economic model developed in Microsoft Excel®.

In response to a number of points for clarification raised by the ERG, the manufacturer further submitted:

- Additional exploratory analyses related to the model time horizon and results for alternative time horizons of 2 and 8 months.
- An additional scenario where treatment is assumed to be maintained for at least 2 years for patients at risk of relapse, based on recommendations from NICE Clinical Guideline CG 90.
- Further analyses of the EQ-5D data from the REVIVE trial.
- Additional clarification regarding the 'switch cost' applied when patients switched therapy and the distributional assumptions and data used for parameters in the probabilistic sensitivity analysis (PSA).

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

Section 7.1 of the MS focusses on identifying evidence specifically on the cost-effectiveness of vortioxetine. The search strategies were briefly described in the main body of the submission, and full details were provided in an appendix. The MS did not attempt to formally identify published evidence on the cost-effectiveness of other antidepressants. However, the manufacturer made reference to the economic model used in NICE CG90 to order to provide a rationale for any differences in model structure proposed within the manufacturer's de-novo analysis.

5.1.1 Searches

A number of databases were searched, including MEDLINE, EMBASE and the Cochrane Database of Systematic Reviews (CDSR). Reporting of search strategies was generally appropriate. An appropriate economics study-design filter was used in the MEDLINE and EMBASE searches.

The search strategy undertaken for the specific aim of the review presented by the MS is considered appropriate by the ERG.

5.1.2 Inclusion/exclusion criteria used for study selection

The inclusion/exclusion criteria used for study selection can be found on page 141 of the MS and are reproduced below:

- Population: adults with major depressive disorder or experiencing major depressive episodes.
- Intervention: vortioxetine. Evaluations without a vortioxetine-based regimen were excluded.
- Outcomes: results from economic evaluations (including cost-utility, cost-benefit, cost-effectiveness, cost-consequence and cost-minimisation analyses).
- Study designs: economic evaluations and reviews of economic evaluations (to source the original studies).

The ERG considers these criteria to be appropriate to identify existing published evidence specifically on the cost-effectiveness of vortioxetine.

5.1.3 Studies included and excluded in the cost effectiveness review

Twenty-eight potentially relevant studies were identified from the search strategy. However, 26 of these studies were subsequently excluded on the basis of the information reported in the title and abstract. The remaining 2 studies were excluded after a full-text review. Consequently, no previously published studies of the cost-effectiveness of vortioxetine were identified in the MS.

The ERG did not identify any additional published evidence which met the stated inclusion/exclusion criteria. However, the ERG also searched for public documents for the submissions to Pharmaceutical Benefits Advisory Committee (PBAC) in Australia and the Canadian Agency for Drugs and Technologies in Health (CADTH). These submissions were highlighted by the manufacturer in their clarification response to issues raised by the ERG concerning the clinical effectiveness data. Only the public summary document for the PBAC submission was subsequently identified and considered by the ERG.

Only limited details were available in the public summary document for the PBAC assessment and a full review of the submission and approach was not possible. The population considered in the PBAC submission appeared similar to the manufacturer's stated decision problem for NICE, namely "patients who have received and not responded to an initial antidepressant medication or patients who are intolerant of or who have contraindications to other initial antidepressant therapy". The economic submission appears to have been based on a cost-minimisation analysis comparing vortioxetine with desvenlafaxine. Desvenlafaxine is an SNRI based on a synthetic form of the major active metabolite of venlafaxine. The relevance of this study to the NICE decision problem is clearly limited since desvenlafaxine is not commercially available in the UK. However, the ERG considers that the approach and the subsequent responses provided by the Economics Sub-Committee of PBAC were a relevant consideration. The public summary documented stated that:

"By cost-minimising vortioxetine to desvenlafaxine alone, the submission suggested that the price of vortioxetine could be made more commercially viable than a price based on a cost-minimisation listing against venlafaxine or duloxetine which the submission had noted to have been affected by statutory price reductions through generic competition. The ESC considered that this was not reasonable. The economic comparison should have reflected the current range of drugs likely to be displaced by vortioxetine.

The ESC advised that the submission should have used a weighted mean price of the drugs that would actually be displaced by vortioxetine. The ESC noted that the evaluation suggested that this could either be done using a weighted mean of SNRIs, or a weighted mean of all alternative therapies (SSRIs, SNRIs, tricyclic anti-depressants and other anti-depressants)".

5.1.4 Conclusions of the cost effectiveness review

In the absence of any previously published cost effectiveness studies of vortioxetine for the management of MDE, the de-novo analysis in the MS represents the most relevant evidence for the stated decision problem.

The ERG recognises that the manufacturer has identified and subsequently justified differences between their approach and the model used to inform NICE CG90. However, the ERG also considers that a more formal review of existing economic evaluations for other antidepressants would have been helpful in providing further justification for their approach and to assist with validation (i.e. the extent to which the manufacturer's model for comparator treatments is consistent with previous models). Such a review could also have provided a helpful conceptual basis for informing and justifying structural choices and assumptions. This seems particularly pertinent since the single study where a comparison is made (NICE CG90) evaluated the cost-effectiveness of 1st line therapy and did not formally include the decision (and associated costs and consequences) of switching therapy.

The ERG identified a previous published systematic review of methodologies used in cost-effectiveness models for evaluating treatments in major depressive disorder.(34) This review identified 37 studies; 29 of these included pharmacological interventions, 9 of which were UK studies. The review highlighted that the model time horizon and associated structure were important aspects in capturing the costs and consequences of the different treatment phases. These phases include acute treatment, during which time the goal is to resolve symptoms; continuation treatment, during which time therapy is continued to ensure complete resolution of the episode and to prevent relapse; and long-term maintenance, where therapy is continued to prevent the development of a new episode.

The majority of the models reviewed used a decision-tree structure (n=28) and most of these had a time horizon of 6 months to 1-year, which was typically separated into two intervals representing the acute and continuation phases of depression.

The review highlighted that 15 of the studies were based largely on 2 alternative structural approaches based on those initially developed by Francois (35) and Casciano.(36)

- The decision-tree structure developed by Francois (35) consisted of two main pathways. Patients with MDD entered the model in the first path representing primary care. Patients with inadequate response in primary care could subsequently titrate to a higher dose or switch treatment. Patients with insufficient response after titration and/or switching were referred to secondary care, which was represented by the second path. In the secondary care path, patients could have their dose titrated, have their treatment switched, receive adjunctive therapy with another agent, or be hospitalised. The model structure also incorporated the risk of suicide and attempted suicide.
- The decision-tree structure developed by Casciano (36) included additional events following treatment failure due to lack of efficacy including titration to maximum dosage, within-class adjunctive therapy, between-class adjunctive therapy, and treatment switch. The model structure did not incorporate the risk of suicide and attempted suicide.

Several key differences were highlighted in the review between these two modelling approaches. Firstly, the models incorporated different treatment options for patients failing first and second lines of treatment, with the Francois structure also including the option of hospitalisation once all treatment options had failed. Secondly, the Casciano structure assumed that successfully treated patients (defined as a 50% or greater improvement in MADRS or HAMD score) would continue on treatment for 6 months and didn't include the risk of a subsequent relapse. In contrast, the Francois structure incorporated a risk of premature treatment discontinuation as well as the risk of a subsequent relapse.

The review also highlighted the significant variation across studies in the measures of treatment success that were used, with measures based on response and/or remission. The lack of consistency in the measures and the definitions of these measures were also noted by the authors. Response was most commonly defined as a 50% or greater improvement in the MADRS score or the HAMD-17 score. However, one model used three definitions for different levels of response based on MADRS scale: response (greater than 50% improvement from baseline), partial response (25%-50% improvement from baseline) and no response (< 25% improvement from baseline).

The review concluded that there appeared general consensus that values of 7 or less on the HAMD and 10 or less for MADRS were indicative of clinical remission. However, of the 14 models using the MADRS-based definition of remission, only two used a cut-off value of ≤ 10, with the majority of the models using a cut-off value of ≤ 12.

The systematic review of existing model structures is useful in highlighting the variation that exists both in terms of how initial success is determined (i.e. response and/or remission) and the alternative pathways which are considered following an initial successful or unsuccessful treatment. However,

the review provides limited insight into the appropriateness of using alternative approaches to defining success. The ERG considers this an important limitation since inevitably the criteria applied and the timing of this is an important consideration in ensuring that the model structure appropriately characterises the way in which clinical decisions are made as well as the subsequent pathways followed by patients.

Importantly, while the majority of studies appeared to use either response or remission as the determinant of treatment success (such that subsequent decisions and pathways were then conditioned upon this measure), 4 of the 8 UK studies included both measures. The studies by Benedict (2011) (37) and Lennox-Smith (2008) (20) are particularly insightful since these studies include additional pathways for patients who have responded but not yet achieved remission (i.e. responder/non-remitters). Both models assumed that the initial treatment would be continued in responder/non-remitting patients beyond the initial acute period (8 weeks in both models). This approach contrasts with other published models which typically assumed a single continuation rule based on either response or remission at a single time point. The ERG clinical advisors consider that these additional pathways included by Benedict (2011) and Lennox-Smith (2008) appear more reflective of clinical practice than a single continuation rule based on either response or remission.

Although both Benedict (2011) and Lennox-Smith (2008) allowed patients who have responded but not yet achieved remission to continue therapy beyond 8 weeks, they differ structurally in terms of subsequent pathways and timings. Lennox-Smith (2008) incorporates a single additional 8-week period for this group, such that if these patients do not subsequently achieve full remission over the next 8 weeks (i.e. 16 weeks in total) then their initial therapy will be switched at this point. In contrast, the Markov-structure employed by Benedict (2011) assumes that therapy can be continued in this group for up to 40 weeks (i.e. covering the entire 48 week horizon) as long as the initial response is maintained.

The systematic review is helpful in highlighting the significant variation that exists within existing models and the different approaches employed concerning key structural assumptions. However, it is evident that no clear consensus appears to have emerged on the most appropriate structural assumptions for modelling an MDE and alternative treatment strategies. Despite the lack of consensus, the ERG considers that the review provides an important basis to critique the approach undertaken by the manufacturer in their de-novo analysis for vortioxetine. In particular, the following points are important considerations:

1. The time-horizon should be sufficient to capture the different phases of treatment. Several authors have argued that a horizon of 1-year is appropriate to capture the full recommended course of treatment (6 months after remission) and relapses within a year.
2. Heterogeneity based on the severity of the initial episode (or other characteristics such as number and timing of previous episodes) may mean that different model pathways or structures should be considered. For example, several of the existing models assume different pathways for patients being managed in a primary or secondary care setting. The study by Benedict (2011) assumed that patients with a score of ≥ 25 on the HAMD-17 were likely to be referred to mental health specialists in secondary care.
3. There appears to be significant inconsistency amongst existing studies concerning the appropriate definition of initial treatment success and subsequent clinical decisions. While the majority of existing models use either response or remission (variably defined across studies), several of the more recent publications have incorporated both response and remission as outcomes. These studies have incorporated separate pathways for patients achieving remission during the initial acute period as well as for patients who have responded but not yet achieved full remission. Importantly, those studies which have incorporated these additional pathways have been those in which the decision to switch therapy (and associated pathways) has been a central consideration.
4. There also exists significant variation in subsequent pathways for patients not achieving an initial treatment success. While some models assume that patients are subsequently switched onto another antidepressant medication, other models assume a variety of strategies including titration and augmentation, including both pharmacological and non-pharmacological interventions (e.g. psychotherapy, ECT, hospitalisation).

5.2 The manufacturer's economic evaluation compared with the NICE reference case checklist

Table 27 NICE reference case checklist

Attribute	Reference Case	Included in MS	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Comparator(s)	Alternative therapies in the NHS, including those currently regarded as current best practice	Partially	The comparators used do not represent the full set of alternative therapies. 18 pharmacological comparators were included in the scope, which were not included in the cost-effectiveness model. These included augmentation therapy as well as competing antidepressant regimens.
Type of economic evaluation	Cost-effectiveness analysis with full incremental analysis	Yes	Yes. The *de novo* model produced in MS Excel considers pairwise comparisons only, requiring a few simple steps in order to produce full incremental analysis.
Perspective - costs	NHS and Personal Social Services	Yes	The MS also considers a wider societal perspective; where the costs of absenteeism were included through consideration of data relating to reported sick leaves collected as part of PERFORM. These are combined with national average data on wages. A number of wider societal costs are missing and so this is referred to as a partial societal perspective.
Perspective - benefits	All health effects on individuals	Partially	The MS focuses on the achievement of remission in patients, but fails to distinguish between responding non-remitters and non-responding non-remitters. As a result, some health effects may be missed and the structure of the model does not exactly follow the clinical decision problem as discussed in more detail in this section.
Time horizon	Sufficient to capture differences in costs and outcomes	Partially	The ERG considers that the adoption of a 12-month time horizon seems broadly appropriate given the natural history of an MDE for the average patient, but notes that continuation/maintenance therapy for patients with high risk of relapse would require a longer time horizon.
Synthesis of evidence on outcomes	Systematic review	Partially	All searches within systematic reviews are appropriately specified, however, much discussion of the appropriateness of the narrow population being considered is given in sections 3 and 4.
Outcome measure	QALYs	Yes	
Health states for QALY measurement	Described using a standardised and validated instrument	Yes	Health states were described using the MADRS score, which is a validated instrument for MDD. The MS also attempted to synthesise evidence where HAM-D score was used, using appropriate cross-walk procedures.
Benefit valuation	Time Trade Off or Standard Gamble	Yes	Time Trade Off
Source of preference data	Representative sample of the public	Yes	
Discount rate	3.5% for costs and benefits	Yes	Since the time horizon was 12 months, no discounting was required in the base case. In clarifications, the manufacturer supplied results from a 24 month time horizon, where 3.5% discount rates were applied to costs and benefits.
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	No additional weighting was given to QALYs.
Sensitivity analysis	Probabilistic sensitivity analysis	Partially	Base case results were based on deterministic model, but probabilistic sensitivity analysis was also presented.

5.3 ERG's summary and critique of manufacturer's submitted economic evaluation

5.3.1 Population

As stated in Section 3.1 above, the patient population considered within the manufacturer's decision problem is restricted to a subset of the licensed patient population, namely:

"*adult patients with moderate-to-severe MDD who are experiencing an MDE, who have responded inadequately in terms of efficacy or tolerability to initial antidepressant treatment, and who require and want to switch to alternative antidepressant*" (MS, p14).

This patient group is referred to in the MS as the "switch population". No additional patient sub-groups were considered within the MS.

As previously highlighted in Section 3, the restriction to the switch population appears consistent with the "Other considerations" section specified in the NICE scope. However, the ERG also considers that that the definition of the "switch population" proposed by the manufacturer imposes a further restriction, in that it focuses entirely on the population in whom a single initial antidepressant medication has failed due to lack of efficacy or tolerability. However, clearly there will also be patients in whom a 2nd or a 3rd line of treatment may have failed and who are not directly considered by the manufacturer's analysis.

Inevitably, the current pathway for patients with MDE potentially involves a series of sequential steps and treatment options with the number of steps and options determined by the success of each prior step or treatment option. The ERG considers that the appropriate population and potential position of vortioxetine within current pathways should have been more formally considered by the manufacturer, based on a broader consideration of the evidence based for vortioxetine and other comparators, rather than restricting the decision population and evidence base from the outset. Furthermore, by only considering the use of vortioxetine as a second line treatment, the MS is not sufficiently flexible to establish the optimal position of vortioxetine within the existing treatment pathway. The ERG considers that the restrictive presentation of the decision problem in the MS prevents a full and appropriate consideration of the optimal position of vortioxetine within an existing sequence of treatment options for the management of patients with an MDE.

5.3.2 Model structure

The MS presents a decision model that evaluates the progression of a single MDE. The model is based on treatment success defined in terms of remission at 8-weeks. The model considers three

stages of disease progression: the acute phase (2 months duration), a maintenance phase (6 months duration), and a recovery phase (4 months duration). Consequently the time horizon of the model is 12 months.

A schematic diagram of the model is presented in Figure 13 below. The model combines a decision-tree structure with a separate Markov process depending upon particular pathways that a patient may follow in the decision-tree.

The initial decision tree-structure is common to all patients during the initial acute phase period (0-2 months). During this period patients may achieve remission or they may withdraw from their current therapy due to short-term side-effects or failure to achieve remission. Patients achieving remission in the initial acute phase period subsequently continue within the main decision-tree structure. Between months 2 and 8, these patients subsequently follow additional pathways (or branches) covering the maintenance phase. During the maintenance phase, patients are subsequently assumed to stay on treatment (i.e. sustained remission) or they may stop treatment due to an adverse event or subsequent relapse. If patients sustain remission during the maintenance period, they enter the final part of the decision tree structure representing the recovery phase which covers the final 4 months of the total 12-month time horizon. Importantly, during the recovery period, the therapy is assumed to be discontinued and an assumption is made that patients are no longer at risk of relapse or recurrence.

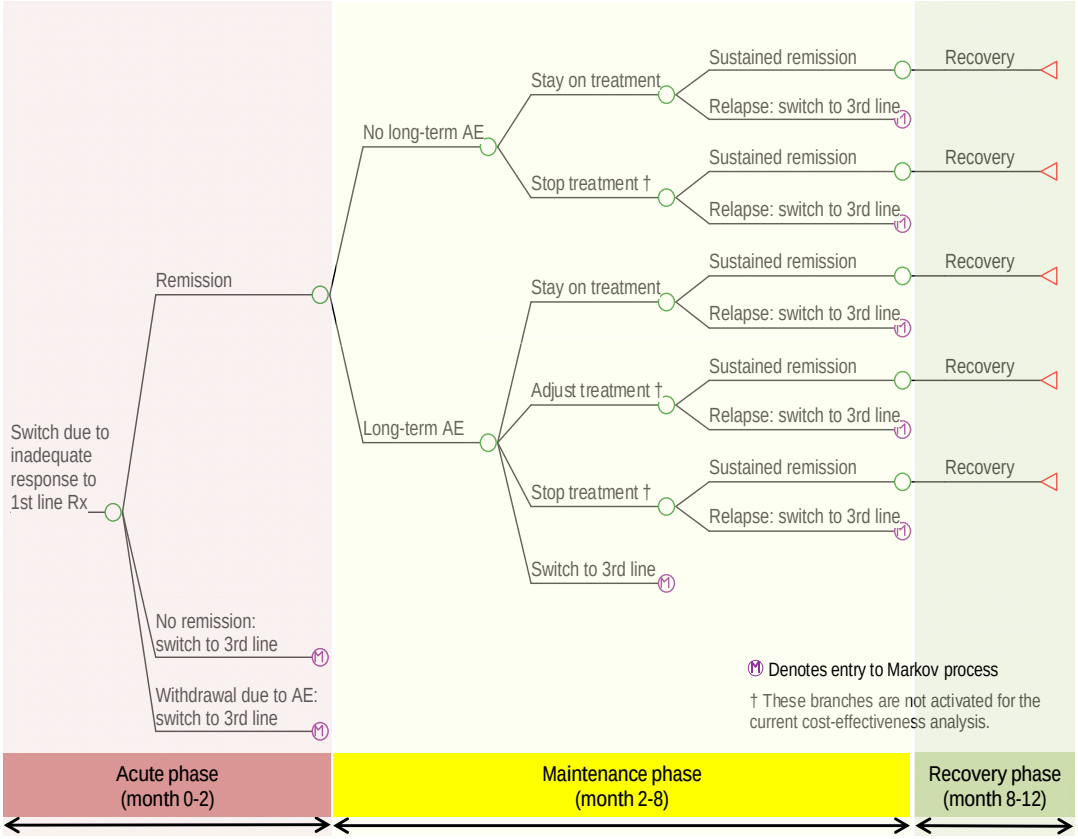


Figure 13 Model schematic diagram (MS Figure B21)

Consequently, the full decision tree is only followed for the entire 12 month period in patients who achieve remission during the acute phase and subsequently maintain this remission during the entire maintenance period. In contrast, all other patients enter into a separate Markov process, with a 2-monthly cycle length. The time at which they enter the Markov process depends upon the phase that the initial therapy is assumed to be stopped (i.e. acute vs maintenance period).

The separate Markov process is then used to model subsequent lines of therapy, with the model allowing up to a maximum of 3 additional lines of treatment (i.e. covering 3rd – 5th lines of therapy, since the initial therapy in the switching population is already the 2nd line of treatment in the overall management pathway).

Figure 14 below illustrates the Markov model schematic in the MS. The Markov model aims to track subsequent patient prognosis in terms of their remission and/or recovery status following additional lines of treatment.

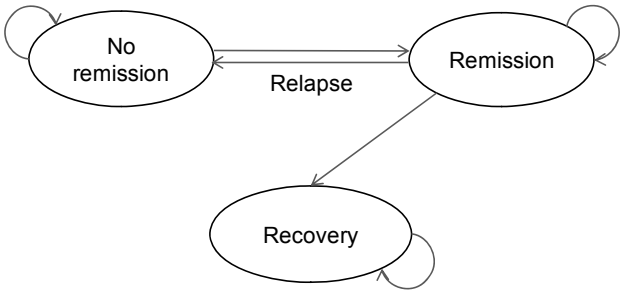


Figure 14 Markov state transition diagram for modelled events following switch to third line (MS Figure B22)

The Markov model uses a 2-month cycle and following every cycle in the no remission state (2 months) an additional switch and further line of treatment is assumed. This means 2 months after initiating a subsequent treatment line patients either achieve remission or switch to an additional line of treatment again.

The manufacturer's structure broadly characterises the progression of a treated MDD episode making reference to a diagrammatic representation (reproduced in Figure 15 below) adapted from Bakish (3).

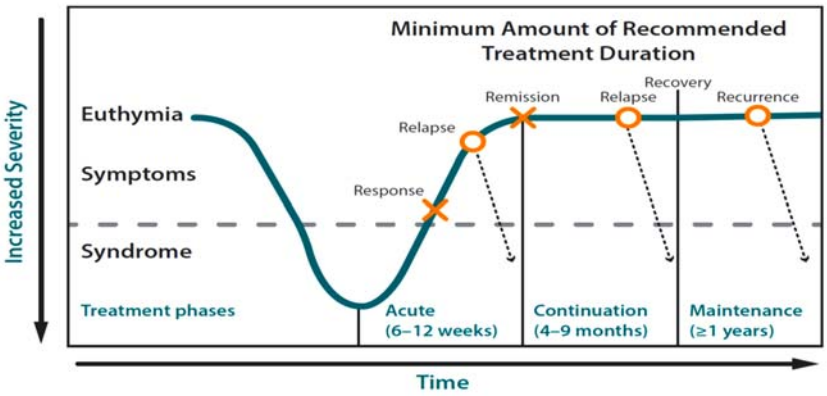


Figure 15 Diagrammatic representation of the course of an MDD episode (MS Figure A3, adapted from Bakish)

This characterisation distinguishes three treatment phases:

1. Acute phase (up to 6 to 8 weeks): the goal is to elicit a response and decrease symptoms to a non-pathologic level
2. Continuation phase (lasting 4 to 9 months): the goal is to maintain the improvements, resolve remaining symptoms and functional impairments and prevent relapse