HIGHLY CONFIDENTIAL

Single Technology Appraisal

Esketamine for treatment-resistant depression [ID1414]

Committee Papers

HIGHLY CONFIDENTIAL

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Esketamine for treatment-resistant depression [ID1414]

Contents:

The following documents are made available to consultees and commentators:

1. **Response to consultee, commentator and public comments on the second Appraisal Consultation Document (ACD2)**

2. **Comments on the ACD2 from Janssen**

3. **Consultee and commentator comments on the ACD2 from**:
   a. SANE
   b. British Association for Psychopharmacology
   c. Royal College of Psychiatrists

4. **Comments on the ACD2 received through the NICE website**

5. **Evidence Review Group documents prepared by Kleijnen Systematic Reviews (KSR)**
   a. ERG critique of company comments on the ACD
   b. ERG response to the additional factual inaccuracies raised by the company

6. **Company response to the critique prepared by Kleijnen Systematic Reviews (KSR)**

7. **Company submission addendum following topic pause after ACM3 from Janssen**

8. **Evidence Review Group critique of company addendum prepared by Kleijnen Systematic Reviews (KSR)**

9. **Comments on the company addendum from NHS England and NHS Improvement**
   a. Mr Peter Pratt – commissioning expert, nominated by NHS E&I
   b. Professor Gary McVeigh – clinical adviser to NHS E&I

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

Esketamine for treating treatment-resistant depression

Single Technology Appraisal

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

Type of stakeholder:

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

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

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

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

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

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
1	Consultee (professional group)	British Association for Psychopharmacology	From NICE: "But it is unclear how effective esketamine is because of the way the trials were done." In discussing the trials, the committee noted that TRANSFORM 1 and TRANSFORM 3 did not show significant results. BAP response: We note that nasal esketamine has been deemed an effective medication by 2 major regulatory authorities (FDA and EMEA) which have approved it for patients with Treatment Resistant Depression. It is important to note that Transform -2 allowed flexible dosing of esketamine, which is in line with the relevant Summary of Prescribing Information. However, the regulatory approval was based on data from 2 positive phase 3 studies (studies 3002, Transform 2 and 3003, Sustain 1), as well as supportive data from additional phase 2 and 3 studies, demonstrating consistent efficacy of esketamine and evidence supporting long-term safety (study 3004). To approve a new antidepressant, our understanding is that regulatory authorities generally require two positive, short-term, adequate and well-controlled studies to meet the (regulatory) standard for substantial evidence of effectiveness. Randomized withdrawal studies are typically conducted after approval to support an additional maintenance claim. For esketamine, however, regulatory authorities required both short- and long-term data in the initial application due to the novelty of the product. It should also be noted that, when the results of TRANSFORM 1 and 2 are pooled, the results do indicate a significant effect. Therefore, we would conclude that for adults aged 18-65 years significance was demonstrated.	Comments noted. The committee were aware of trials used to support the regulatory authority submissions. How the data from these trials is used in the model is considered in section 3.8 of the FAD. The committee further considered the trial evidence as it relates to the treatment effect and subgroups in sections 3.10 and 3.11 of the FAD.
2	Consultee (professional group)	British Association for Psychopharmacology	From NICE: "The response and remission evidence from TRANSFORM-2 should be considered with caution because of the short duration of the trial" The committee took into account a Consultee who stated this time period had little bearing on the treatment for depression, though no evidence is given for this statement. BAP response: We do not understand this point raised by the committee. It is unclear why the committee should choose to reference	Comments noted. The committee considerations about the duration of the trial are

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			the NICE guideline on depression and antidepressant treatment, given that esketamine is a different class of drug, with a different mode of action-and effects on depression scores, with separation from placebo early on in treatment. If anything, effects at 4 weeks are likely to be an underestimate of overall effect, due to the fact that response and remission rates would not likely decrease in value both in the interventional and the control arm (but would only likely increase) with a longer observation period. In addition, study 3003 demonstrated a statistically significant long-term effect of esketamine plus oral antidepressant in maintaining a state of remission or response when compared against oral antidepressant alone. An additional point to make about the duration of trials is that 28 days may seem short for trials of antidepressants that are typically taken for many months, but this is the internationally agreed frame for a licensing trials because of the ethical difficulty of leaving people on placebo for longer periods. Therefore, it is inevitable that prelicensing longer term studies are open label, and that post licensing studies are used to clarify longer effects. We appreciate the NICE committee's frustration with this, but it is the reality of research into all new drug treatments for depression and it does not seem reasonable to withhold a drug from the widespread use that is necessary to obtain the long term information. The overall evidence clearly does not support the logic implied in the comment, that the results of a 4-week duration trial has no impact on real-world clinical significance.	outlined in section 3.14 of the FAD. Section 3.19 of the FAD also details the importance of this data for economic modelling. The committee noted during the appraisal process that the ethical difficulty of leaving people on placebo for longer periods did not apply in this case because of the active comparators in the placebo arm.
3	Consultee (professional group)	British Association for Psychopharmacology	From NICE: The TRANSFORM-2 study is not powered to detect difference in effect between treatment arms so could show a false positive result BAP response: The committee rightfully point out that in this trial a higher than normal placebo response than would be expected was seen. They also highlight potential regression to the mean. Both of these aspects of trial design mean that any placebo/drug difference would be minimised. Therefore, (in our view) the correct interpretation is that these data probably underestimate the treatment effect. Furthermore, the committee also point to the short, 4-week duration-another factor that would minimise the drug/placebo difference. The fact that the initial power calculation was based on a higher estimated difference does not seem relevant here, as the effect size is reported in the trial.	Comments noted. The committee noted the difficulty of interpretation of results with high placebo response in section 3.13 of the FAD.
4	Consultee (professional group)	British Association for Psychopharmacology	From Nice: "Withdrawal effects are difficult to distinguish from symptoms of depression" The report quotes a consultee who queries whether withdrawal from esketamine could confound relapse rates. No evidence is given. BAP response:	Comments noted. The relevant paragraph has been updated

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			The evidence that exists with regards to withdrawal effects is from people who misuse ketamine, and of the two clinical reports in the literature (both from reviews), less than 50% of ketamine abusers developed withdrawal symptoms. These people were using ketamine daily, at doses up to 9g-far in excess to that used in the esketamine trials. Furthermore, there is very little information on cardinal features of withdrawal in these reports. In Study TRD3003, although there was a high number of relapses in the first month in those switched to placebo nasal spray, it is unlikely that a pharmacologic withdrawal effect contributed given that the decrease in esketamine plasma concentrations is rapid for the initial 2 to 4 hours and more gradual thereafter (with a mean terminal half-life, 7-12 hours), with steady state never reached with intermittent dosing. Moreover, this high rate of early relapse is similar to that observed after cessation of electroconvulsive therapy. There are no known rebound effects after electroconvulsive therapy discontinuation. The high rates of early relapse after esketamine discontinuation and those observed by Rush et al for patients in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study at level 3 or 4 (i.e., who had failed 2 and 3 prior antidepressant treatments, respectively) more likely reflect a greater vulnerability to relapse among patients with TRD during maintenance treatment with an antidepressant alone. The FDA report states that “Acute esketamine withdrawal is likely not a factor, as dosing is infrequent during the maintenance phase.” It is physiologically implausible for such infrequent dosing to cause a withdrawal syndrome. Therefore, we conclude that there is nothing evident to us to suggest a withdrawal syndrome. Furthermore, no evidence is presented to show that items on the withdrawal checklist correlate with those of the MADRS, and the trial authors themselves state, “No evidence of a distinct withdrawal syndrome was observed during the 2 weeks after cessation of esketamine nasal spray as assessed by the 20 item Physician Withdrawal checklist”.	in the FAD. The committee discuss this issue in section 3.15.
5	Consultee (professional group)	British Association for Psychopharmacology	From NICE: The differences in relapse rate in the SUSTAIN-1 trial data should be considered with caution BAP response: This argument was initially put forward in a comment on Lancet Psychiatry (Lancet Psychiatry. 2019; 6: 977-979) and we note has since been addressed by the company (Lancet Psychiatry VOLUME 7, ISSUE 3, P232-235, MARCH 01, 2020): neither the company nor the FDA (after site inspection) found any reason to exclude data from the site in Poland which is the subject of the author’s comment. Nonetheless, a sensitivity analysis was performed excluding this site and using a statistical method appropriate for time to event data. Statistical significance was maintained (log-rank test $p<0.05$) and the results remain consistent with the primary efficacy analysis. It is puzzling why the committee should continue to discuss this point.	Comments noted. The committee considered it appropriate to refer to the regulatory agency decision about data from the site in Poland. Therefore, this section of the ACD2 has been removed in the FAD.
6	Consultee (professional group)	British Association for Psychopharmacology	From NICE: Healthcare resource use costs should be made equal across both arms in the current model. BAP response: • We feel that it is clinically more appropriate to assume that healthcare resource use and hence medical costs is/are higher when patients with TRD are symptomatic compared to when patients are in remission than to assume (as NICE appears to) that	Comments noted. The scenario for healthcare resource use

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			healthcare resource use and medical costs are the same, independent of clinical effectiveness of the treatment received and the health state of patients.	with equal costs between arms was based on uncertainty of clinical events and implausibility of costs in the MDE health state. Section 3.32 of the FAD has been updated to explore an alternative source of healthcare resource use.
7	Consultee (professional group)	British Association for Psychopharmacology	From NICE: The effect of subsequent treatments is underestimated, and the ERG's adjustment is more plausible BAP response o It is important to highlight that the majority of patients who have not responded to many antidepressant treatments have very poor outcomes in the long-term o This is about the total cohort of patients with TRD, and not a subgroup of patients who have received very intense treatment in an extremely specialised hospital setting (which has since closed) and were doing well after their intense treatment as shown in the Fekadu and Wooderson study	Comments noted. Section 3.23 of the FAD highlights the uncertainty with long-term outcomes in the model.
-	Consultee (company)	Janssen	Please note the full response can be found in the committee papers, only summary responses are included here. Executive Summary: Janssen welcomes and thanks NICE for the opportunity to comment on the second Appraisal Consultation Document (ACD 2) for esketamine for treatment resistant depression [ID1414]. Overall, Janssen is disappointed with the decision not to recommend esketamine nasal spray (ESK-NS) for treatment-resistant depression (TRD) in ACD 2, despite the Committee's recognition of the unmet need for patients with TRD and the impact this condition has on patients, families and carers. As noted by the NICE Committee and a number of consultees, TRD is a seriously debilitating, potentially life-threatening condition. With each treatment failure, TRD disease morbidity increases, with reduced quality of life, increased costs and poorer outcomes observed (1). Unipolar major depression is projected to be the leading cause of disease burden by the year 2030 worldwide (2), with much of the overall burden of major depression falling on those with TRD (3). Newly licensed treatments which enable the optimal treatment of TRD are urgently needed given the limited innovation in the disease area for the last 30 years. Mental health represents up to 23% of the total burden of ill health in the UK but only 11% of NHS England's budget (4). Despite the Health and Social Care Act 2012 calling for 'parity of esteem' between mental health and physical health, large inequities remain. This is especially	Company consultation comments have not been responded to as they have been superseded by the company's appraisal committee 4 submission.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			pertinent given the unprecedented COVID-19 pandemic, with increased levels of mental health support and investment required. Janssen therefore remains fully committed to addressing the Committee's concerns raised in ACD 2 and ensuring that patients with TRD and healthcare professionals have the option to access ESK-NS for TRD. This is especially important given the lack of innovative treatment options and inequities that persist in mental health compared to physical health conditions. We note that the Committee's concerns in the ACD fall into 2 key areas: questions around the clinical effectiveness of ESK-NS and the economic model. We understand that some of the clinical concerns have informed the Committee's conclusions regarding the assumptions made in the economic model. Our response therefore addresses the clinical concerns raised by the Committee first, before addressing the economic modelling issues, and is summarised below. In addition, we are aware of the Committee's points in the ACD around the positioning of ESK-NS and some of the Committee's uncertainty regarding the clinical and economic evidence. We have therefore conducted a subgroup analysis and provided additional scenarios for the Committee's consideration of ESK-NS in a later line position in the treatment pathway, i.e., used at 4th line after patients failed 3 or more oral antidepressant (OAD) treatments. This cohort of patients are more resistant and difficult to treat, having failed 3 lines of OAD (5) and the clinical evidence suggests even greater relative efficacy and cost effectiveness of ESK-NS. This is where the higher unmet need in TRD is and therefore ESK-NS has the potential to provide an even more valuable treatment option versus currently available options. In terms of addressing the Committee's clinical concerns, the EMA, the FDA and other regulators around the world have all concluded that ESK-NS is a clinically effective option and has an appropriate risk-benefit profile for the treatment of TRD. We note that many of these concerns have been examined by the regulators previously. It is important to provide the context in which regulators have considered the ESK-NS trials, which includes some unique challenges in conducting clinical research in mental health, which the Committee may also wish to consider. It must be noted that developing new medicines for the treatment of Central Nervous System diseases is challenging, with a 15% overall probability of success in clinical trials (6). Approximately 50% of short-term, randomised, controlled trials for approved antidepressants may still fail to show a statistically significant effect (7), primarily due to high placebo responses, which are particularly prominent in mental health trials as opposed to somatic medicine. Despite these challenges, ESK-NS demonstrated superiority to an active comparator in a population with severe disease that did not respond to at least two previous treatments. The issues relating to the clinical data were also debated by the CHMP, who ultimately judged that the short- and long-term efficacy of ESK-NS in patients with TRD had been established. In an effort to provide a submission package that addresses the specific requirements of the evaluation of a new technology in mental health, Janssen sought NICE Early Scientific Advice (in 2013) and modelling advice from NICE PRIMA (in 2018). It is important for the Committee to note in their consideration of the evidence that we have tried to implement as many of the recommendations from the NICE advice as we have been given. Our response intends to address the additional clinical concerns from the Committee and provide reassurance of the clinical benefit that ESK-NS brings, which was also clearly articulated by the patient representative during the second NICE Appraisal Committee meeting.	
1	Consultee (company)	Janssen	**Key point 1: Regulatory authorities assessed the ESK-NS clinical data (TRANSFORM-2 and SUSTAIN-1) and concluded they are robust and demonstrate the clinical value of ESK-NS** • The clinical data for ESK-NS should be considered in the wider context of the unique challenges of conducting clinical trials in this therapeutic area. Regulatory agencies approved ESK-NS having discussed similar clinical points as raised to NICE by a small number of clinical stakeholders.	Company consultation comments have not been responded to

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			• The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcomes for patients with TRD. • TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding. • Response and remission are established and appropriate outcomes for MDD and TRD. • The four-week duration of TRANSFORM-2 is appropriate and is aligned with clinical trial design guidance from the CHMP. • No further conclusions can be drawn about the proportion in response in both study arms if the duration of the trial would have been longer. • The randomised withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results. • Patients with suicidal ideation were not excluded from ESK-NS trials and patients with high risk of suicide were studied in a separate clinical development program for ESK-NS. • A robust risk management plan has been agreed with the MHRA. • Results from an additional long term safety study show no unexpected safety signals.	as they have been superseded by the company's appraisal committee 4 submission.
2	Consultee (company)	Janssen	**Key point 2: The outcomes predicted by the economic model are reflective of the outcomes that patients with TRD experience in the long term and the proportion of patients in MDE health state is appropriate, especially when a revised method for subsequent treatments is incorporated.** • The health states used in the model are appropriate, established and based on previous depression models accepted by NICE. • Results from a targeted literature review of patients with TRD show that long term outcomes of patients with TRD are poor. • There are limited data to inform the long-term outcomes of patients with TRD. The Fekadu study cohort is not appropriate to compare to the outcome of the economic model, primarily because it only included a specific subgroup of patients who were discharged following intense multi-modal inpatient hospital treatment. Instead, wider literature from the targeted literature review provide more appropriate and generalisable data on the long-term outcomes of patients with TRD. • Given the Committee's concerns with the proportion of people in the MDE state, a revised method for subsequent treatments is proposed which reduces the proportion of patients in the MDE state over time. With the revised method for subsequent treatments, the model better represents the long-term outcomes as per the available literature.	Company consultation comments have not been responded to as they have been superseded by the company's appraisal committee 4 submission.
3	Consultee (company)	Janssen	**Key point 3: The use of the base case MDE utility is appropriate, and an alternative approach which addresses the Committee's concerns using an amended criteria for MDE and a different utility value is provided for consideration** • The use of the MDE utility is appropriate as it represents the quality of life of patients with TRD with moderate-severe disease. • Given the Committee's concerns with the different thresholds used in the model and clinical trial programme, we provide a scenario which amends the criteria of the MDE health state. This means an alternative source from a subgroup of moderate to severe patients is used to inform the utility of the MDE health state, rather than using the TRANSFORM-2 baseline MDE utility. • This change in criteria of the MDE health state to represent the moderate to severe MDE would cover the relapse threshold (≥22) used in SUSTAIN-1, and as such addresses the concerns with the different thresholds used in the clinical trial programme and the model.	Company consultation comments have not been responded to as they have been superseded by the company's appraisal committee 4 submission.
4	Consult	Janssen	**Key Point 4: It is appropriate to assume different healthcare resource use costs per health state, which could lead to different**	Company

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
	ee (company)		**medical costs between treatment arms. We propose a sensitivity analysis with reduced cost differences among health states to address the Committee's concerns and include additional costs that may be associated with commissioning ESK-NS.** • It is not appropriate to use SUSTAIN-1 to inform HCRU per treatment arm as it was not designed to collect resource use data, cannot provide any conclusions due to a small number of events, does not consider the full cohort of patients with TRD and is not generalisable to resource use in UK clinical practice. Evidence shows that differential costs per health state are appropriate. This has also been the approach used in previous NICE decision making. • Given the Committee's concerns with the MDE health state, we provide a sensitivity analysis where the costs of the health states are adjusted using the lower bound of the 95% CI costs, resulting in reduced cost differences among health states. • A sensitivity analysis, with revised health state costs, is provided to address the Committee's concerns, such that it is acceptable to have differential costs per health state. • Estimates of the costs of commissioning are incorporated into the model and have very minimal impact on the cost effectiveness.	consultation comments have not been responded to as they have been superseded by the company's appraisal committee 4 submission.
5	Consultee (company)	Janssen	**Key Point 5: Revised post ACD-2 scenarios are provided for consideration (full label population), and in response to clinical consultation feedback to NICE on the proposed later line positioning of ESK-NS, we have conducted analyses for this subgroup of patients.** • Given the comments in the ACD, we propose scenarios with 1 later line positioning (nonresponse to at least 3 prior treatments in the current episode) based on the available data of ESK-NS for the Committee's consideration. • [redacted]	Company consultation comments have not been responded to as they have been superseded by the company's appraisal committee 4 submission.
1	Consultee (professional group)	The Royal College of Psychiatrists	Has all of the relevant evidence been taken into account? Evidence on the efficacy and safety of ketamine should be considered. The neuromodulatory effects of esketamine in the brain are likely to be identical or extremely similar to ketamine at equivalent doses. In this regard, the highest level of impartial evidence is likely to be Cochrane Reviews. These show that evidence in support of ketamine in depression is generally of poor quality, involving small samples, and with efficacy only shown over brief (clinically likely to be irrelevant) time periods. Also, risk of bias was often unclear, due to a lack of reporting. See: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858. CD011612.pub2/full?highlightAbstract=ketamine%7Cketamin https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858. CD011611.pub2/full?highlightAbstract=ketamine%7Cketamin Some academic institutions have patents and other intellectual property regarding ketamine and/or esketamine. This is more a case in the USA. Such conflicts of interest are often not mentioned when members of those institutions give presentations as to the apparent benefits of ketamine and/or esketamine.	Comments noted. The committee was aware of the issue around using evidence for intravenous ketamine from the technical report and did not consider it to be a key issue compared with the other issues raised.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
2	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on Esketamine is likely to be used later in the treatment pathway because it has a higher treatment burden than other treatments** We agree that the position of esketamine in the treatment pathway is initially likely to be at least fourth or fifth line – i.e. after trials of augmentation. We also agree that for some patients this is because of the burden of treatment. Patients may not drive following esketamine treatment until they have had restful sleep. They can return home using public transport when they are fully recovered. However, we consider that this later use also reflects its expense, novelty and association with a drug of abuse, more than the clinical data. Compared with the alternatives it is not obviously less safe. Therefore, particularly once costs come down, and particularly for patients are well supported, it is likely to be used earlier in the pathway. For some of those who are less well supported it may be more appropriate to provide better hospital or volunteer transport than to withhold the medication until later.	Comments noted. These comments were noted during the committee discussion of the probable positioning of esketamine, and this is outlined in section 3.4 of the FAD.
3	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the response and remission evidence from TRANSFORM-2 should be considered with caution because of the short duration of the trial** We agree that 28 days does seem short as a primary end point for trials of antidepressants that are typically taken for many months. However, we do not think it is right to say that this has 'little bearing' on the treatment of depression. This is the internationally agreed time frame for licensing trials because of the ethical difficulty of leaving people on placebo for longer. Uniquely amongst programmes for a new antidepressant, the short term 28-day data in TRANSFORM are supplemented by the high quality data of the 1 year study SUSTAIN 2. Usually, it is lower quality post licensing studies that are used to clarify longer effects. It would not seem reasonable to withhold this drug from widespread use on the basis of a criticism that can be levelled at all other antidepressants which are in current use. 'The committee acknowledged that splitting the data into 2 groups could have inflated the differences between arms, particularly because the mean reduction in MADRS was near to the threshold for response in both arms at day 28. So, people could meet the criterion for symptom response in 1 arm but only have minimal differences in MADRS score in the other arm'. We do not agree that splitting the data into 2 groups could have 'inflated the difference'. The fact that the difference in remission, which is based on an absolute threshold level of the MADRS, between the two arms in TRANSFORM 2 (21.5%) is greater than the difference in response (17.3%), which is dependent on change relative to baseline level, effectively disproves the possibility of an inflated effect. We would further make the point that response and remission are entirely conventional, pre-specified, measures. This new concept of a 'threshold in response' does not make sense when the difference in MADRS needed to meet criteria for response will vary for each participant depending on baseline. It is no more right to make decisions based on a NICE-generated post-hoc analysis which suggest that an effect size is 'near to threshold' than it would be to make decisions based on company-generated post hoc analyses which showed big effects.	Comments noted. The committee considerations about the duration of the trial are outlined in section 3.14 of the FAD. Section 3.20 of the FAD also details the importance of this data for economic modelling. The committee noted during the appraisal process that the ethical difficulty of leaving people on placebo for longer periods did not apply in this case because of the

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
				active comparators in the placebo arm. Section 3.14 maintains that dichotomisation of data can result in changes in interpretation with further evidence from the 3 or more subgroup.
4	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the TRANSFORM-2 study is not powered to detect difference in effect between treatment arms so could show a false positive result** We do not understand why universally accepted standards for accepting a difference between two arms of a trial, are described as potentially a 'false positive'. It is of course possible that any result could be a 'false positive', but this is why we have accepted norms of statistical significance. The language here seems inappropriate. One would not accept a comment from a company which asserted that their non-significant result was potentially a false negative because a study was underpowered! The powering of the study is based on the number of patients to detect a difference assuming a specific degree of variance. It is possible that the difference was statistically significant despite the smaller than estimated effect size because, even though the difference was smaller, the degree of variance was lower.	Comments noted. The committee further considered the trial evidence as it relates to the treatment effect and subgroups in sections 3.5, 3.10 and 3.11 of the FAD.
5	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on withdrawal effects are difficult to distinguish from symptoms of depression** We agree that it is difficult to conclusively disprove that a new symptom arose because of stopping the drug rather than because relapse. However, the pattern of new symptoms provide important evidence as to which was happening and this does not appear to have been considered. In SUSTAIN 2 (Wajs et al 2020 Supplementary 5) the following effects were common (all >20% in the second week after cessation): insomnia, anxiety-nervousness, dysphoric mood-depression, fatigue – lethargy – lack of concentration, irritability, difficulty in concentration. These are all symptoms of major depressive disorder. By contrast the following symptoms were much less common: loss of appetite, nausea -vomiting, diarrhoea, poor coordination, sweating, tremulousness, dizziness-lightheadedness, headache, muscle stiffness, weakness, increased acuity sound smell touch, paraesthesias, depersonalisation-derealisation. With the exception of loss of appetite, these are not features of major depressive disorder. The dominant problem is therefore more likely to be relapse in depression rather than new symptoms occurring due to a change in physiology induced by the drug. Increased feeling of hopelessness on withdrawal are an important problem, but are much more likely to be due to relapse in depression rather than being caused by the drug. The short acting nature of the drug means that if it did induce some sort of change of physiology which caused withdrawal symptoms,	Comments noted. This has been updated in section 3.15 of the FAD.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			then these effects would be expected to occur between each weekly dose (thereby undermining its beneficial effect), rather than after the end of a course. This was not observed. For these reasons, we think the results of SUSTAIN 1 should be taken at face value. The main implication of SUSTAIN 1 is that the drug needs to be taken continuously to prevent relapse. It undermines the company's assertion, made on the basis of much less direct evidence, that relapse will not occur if it is withdrawn later.	
6	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the differences in relapse rate in the SUSTAIN-1 trial data should be considered with caution** There seems to be a disparity between the conclusion – that the results of SUSTAIN 1 should be treated with caution – and the text which follows, all of which seems to point to reasons why the data of an outlier should not be excluded. The choice of language here seems inappropriate.	Comments noted. The committee considered it appropriate to refer to the regulatory agency decision about data from the site in Poland. Therefore, this section of the ACD2 has been removed in the FAD.
7	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the evidence for esketamine is limited in its generalisability to the NHS** **Severity** We agree that the trial data are limited in the degree of generalisability to populations that are more severe, but do not think this is a strong argument against adoption. Current practice is to use the same antidepressants in people with depression of all severities. The choice of antidepressants at different points on the treatment pathway is determined by side effect profile rather than by different antidepressants having different efficacy in different severities. Comorbidities The poor generalisability associated with the exclusion of patients with comorbidities is also relevant, both to safety and efficacy. Most psychiatric disorders, are associated with depression; and each will sustain and fuel the other. This is a contributory factor to high rates of prescribing of antidepressants in the population. We think the appropriate way to manage the risks of prescribing in patients with comorbid illness is through good phase 4 studies following adoption, rather than by withholding the drug from people with 'pure' resistant depression because it might be used in people with complicating comorbidities. This can work in unexpected ways. For example, there are data from multiple studies suggesting that ketamine can be of benefit in reducing substance misuse. Clearly, however, there are also risks in people who are vulnerable to developing addiction, as reflected in the datasheet.	Comments noted. The committee considerations around generalisability of the trials are outlined in section 3.16 of the FAD. The committee did not use this as an argument against adoption but instead as an increase in uncertainty for costs and

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
				benefits, particularly with the interaction of esketamine's expected use later in the treatment pathway (section 3.4 of the FAD).
8	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on it is not appropriate to adjust the efficacy estimates of the placebo arm in the trials** Whilst we agree with the company's assessment of the influences on the placebo effect, we agree with the committee that the sort of post-hoc adjustment which the company applied was not appropriate.	Comments noted. Further discussion of the adjustment for placebo effects is outlined in section 3.13 of the FAD.
9	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on safety must be considered when administering and monitoring esketamine** We agree with the committee that a registry is required. Further, we consider that such a registry should be interrogatable. Otherwise, those wishing to prescribe other rapidly acting antidepressants (eg IV ketamine) cannot be sure whether an individual is additionally taking esketamine nasal spray and is 'topping up'. This would be a first step on the way to the use of systems such as Safescript (now mandatory in Australia) and Drug Prescribing Monitoring Programmes (as used in every state of the US) which require that, before prescribing, doctors intending to prescribe certain scheduled drugs must interrogate databases to ascertain existing and previous scheduled drug use. We agree with the committee and the regulators that the signal is not, at present, strong enough to justify withholding the drug from the larger number who may benefit. An interrogatable registry will help in tracking the extent of suicidal behaviour associated with relapse or non-response to esketamine. This is a phase 4 task.	Comments noted. Discussion of the risk management plan and additional long-term safety results is in section 3.18 of the FAD.
10	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on economic model** 3.17 – 3.19 The company's economic model does not reflect the course of the disease We agree that there are 'minimal long-term outcome data for people with treatment-resistant depression' to inform modelling. The higher cost of the drug makes this more important than it is for other cheaper oral antidepressants.	Comments noted. The committee conclusions on long-term outcomes in TRD are outlined in section 3.23 of the FAD.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
11	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the effect of subsequent treatments is underestimated and the ERG's adjustment is more plausible** We agree with the committee that clinical practice would not be that 3 treatments would be attempted within 12 weeks as each successive treatment failed. Cycling between treatment takes much longer than this.	Comments noted. The committee conclusions on subsequent treatment modelling are outlined in section 3.22 of the FAD.
12	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on the cost of a course of esketamine treatment may be underestimated** The committee is concerned about variations in the dose and frequency of treatment. This data already exists. The company's data, as submitted to the FDA, shows that a higher proportion of those who remit but do not respond take maintenance esketamine at the shorter, weekly maintenance interval (69%) than those who remit (34%). In other words, those who respond less well take it more frequently.	Comments noted. The committee were aware of this data from the company submission. Further points are discussed in section 3.31 of the FAD.
13	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on A 1 to 2 ratio of nurses to patients is an appropriate resource cost during post-administration monitoring** We disagree slightly with the committee here. Based on the experience of the 5 UK centres which administer IV ketamine - for which the recovery time and requirements are likely to be similar if not slightly higher than for nasal esketamine - we consider that a 1 to 3 ratio more accurately reflects the need for healthcare staff supervision. Post treatment observation can be done by a healthcare assistant and, depending on the layout of the clinic needs only to be intermittent rather than continuous. It does not require a qualified nurse. We agree with the clinical expert that the staffing need will change as clinics develop experience and efficiency of procedures. A typical ECT department would be able to start by treating esketamine patients at the end of their twice weekly ECT lists, thus avoiding employment of new staff until numbers justified a new bespoke clinical session. Fairly quickly, a single nurse and healthcare assistant can run a clinic with 3 concurrent patients each of which will be in clinic for about 2 hours in total. In a clinic which has the beds/chairs to manage 3 simultaneous patients, two staff would be comfortably able to treat 6 patients in a session, including time for recording notes and, depending on its complexity, completing the registry. It is important to note that, as with directed observed administration of other CDs, a doctor does not need to be immediately present for the treatment.	Comments noted. The committee heard from NHS commissioning experts that the company plan to implement esketamine in ECT suites may not be feasible (see section 3.34). The committee conclusions on the nursing and monitoring cists are discussed in section 3.35 of the FAD.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
14	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on significant investment will be needed to use esketamine in the NHS, but costs are difficult to quantify** Based on our experience, we think the only physical infrastructure likely to be required in an ECT suite is a Controlled Drug cabinet. In other settings it may also be necessary to purchase suitable comfortable chairs. The processes for transporting drugs to the ECT exist already and are part of routine hospital transport systems so this does not incur new costs. The arrangements for disposal of used devices consist of putting a bespoke bin (like a large blue sharps bin) in the department which, when full, is transferred back to pharmacy for formal disposal of the remnants of the devices. This uses existing transport arrangements and again is low cost. Training: The procedure is not complex, training materials are provided by the company and this could be accomplished within existing allocation of training time.	Comments noted. Discussion of these implementation costs are outlined in sections 3.34 and 3.35 of the FAD.
15	Consultee (professional group)	The Royal College of Psychiatrists	**Comment on it will take time and resource use for esketamine to become part of clinical practice** We agree that esketamine is potentially disruptive to existing practice but observe that this may be a good thing. For example, patients with resistant depression commonly find that they become disillusioned with CMHT services because, however good the support, their condition does not change (by definition). When they have a treatment which abruptly helps, their care rapidly aligns with the service which provides it. In our experience, this commonly then results in the CMHT wishing to discharge the patient. The service providing esketamine then finds itself with a rapidly increasing caseload of patients who, if they relapse, are potentially at high risk. One way of managing this risk is to have shared care with the CMHT, but this duplicates effort and can seem pointless to the patient. A better solution may be to draw the resource into the new service from the old. This sort of disruption is to be welcomed – but, like all disruption, may initially be unpopular. We agree that esketamine services should not be confined to ECT services and that community settings would be suitable. However, the infrastructure – a clinic with comfortable chairs, separated by curtains, which is suitable for administration and recovery - is common in many NHS settings. We also agree that the reality of NHS processes is such that the lead times of 6-12 months for implementation quoted are realistic. However, this is driven by institutional barriers to introducing new technologies. Because the 'technology' is very simple, private clinics will be much quicker in set up. In conclusion, we would not describe the costs of setting up a clinic as 'substantial'. The staff running costs could reasonably be estimated as a third of a session of a band 6 and a band 3 nurse – about £20 per patient per treatment.	Comments noted. The treatment setting is discussed in sections 3.34 and 3.35 of the FAD.
1	Consultee (patient group)	SANE	We cannot comment on whether all of the relevant evidence has been taken into account, or on whether the summaries of clinical and cost effectiveness are reasonable interpretations of the evidence. As regards whether the provisional recommendations are sound and a suitable basis for guidance to the NHS, we welcome the committee's conclusions that treatment-resistant depression has a negative effect on people, their families and carers; that the effectiveness of current treatments for the condition is limited; and that there is an unmet need for new treatment options. We hope that, given these conclusions, further work will enable the committee to recommend esketamine for use in the NHS. In undertaking this work, we hope the committee will give due weight to the evidence given by patient experts.	Comments noted. The committee considered the unmet need in treatment depression in section 3.3 of the FAD. NICE

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			We reiterate the points we made in our submission of July 2019 and our comments in response to the appraisal consultation document issued in January 2020. Since we made those submissions, we have seen the effects on mental health of the Covid-19 pandemic. People with depression contacting us on our helpline have told us that they are being severely affected by the pandemic, experiencing a deterioration in their mental health and expecting a further toll with the continuation of restrictions and mental health services not always easy to access. For those with treatment-resistant depression, we fear there could be a much more deleterious effect, putting people at risk of self-harm and suicide. We believe this makes it all the more important and urgent that there be the most effective treatment response for those living with treatment resistant depression.	has continued its appraisal programme work during the COVID-19 pandemic to ensure guidance is delivered as soon as possible.
2	Consultee (patient group)	SANE	We know that for those living with treatment-resistant depression, there is a loss of hope that it can improve, or that any treatments might be helpful or effective. The expert patient evidence in the appraisal consultation document highlights this, and how the feelings of hopelessness increase when multiple courses of treatment do not work. We believe the fact that esketamine can have an effect within 24 to 48 hours of being administered, potentially saving patients the weeks or months of uncertainty that can be experienced with other anti-depressants, is critically important in offering hope to those for whom other treatments have not proved effective, and in seeking to alleviate the negative effect of the condition on patients and carers.	Comments noted. The hopelessness associated with treatment resistance is discussed in section 3.1 of the FAD.
3	Consultee (patient group)	SANE	We are pleased that the committee concluded that the biological mechanism of esketamine could be innovative. Patients are having to rely on medications that are over 30 years old, which can have unpleasant side effects and do not work for everyone. Allowing esketamine as an additional treatment would offer the possibility of relief from suffering, and widen patient choice.	Comments noted. The innovative nature of the biological mechanism is discussed in section 3.39 of the FAD.
1	Commentator (web comments)	-	Has all of the relevant evidence been taken into account? Patients suffering from treatment resistant major depression use more healthcare resources in terms of hospital admissions, GP consultations and psychological treatments than patients who have responded to treatment and recovered. It is vitally important to have another helpful and useful treatment other than ECT that patient can access. E Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? There are Clinical Treatment Teams as well as ECT Suites in some areas, which can form the spine for setting up the clinics. They have resuscitation and monitoring equipment as well as appropriately trained professionals. This consideration will give a better interpretation of cost-effectiveness.	Comments noted. Please see the updated guidance document (FAD) for further information.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			Are the recommendations sound and a suitable basis for guidance to the NHS? The conclusions made by NICE at this stage of lack of cost-effectiveness is misjudged. It is in the interest of the NHS and our patients to grant approval for use in combination with an SSRI or SNRI as a third line choice. The goal must be to improve the patient's quality of life and where possible a return to gainful employment.	
1	Commentator (web comments)	-	General Comment: As someone with treatment resistant depression I need more options. I have frequent suicidal thoughts and attempts. I have heard first-hand accounts of people in the US who have benefited from esketamine and I want that opportunity to be relieved from my depression. Depression affects so many people in my life, not just me. My husband, family and friends. SSRIs are limited and almost impossible to come off. We need other options. The money that this will cost may save thousands of lives. If this was a new cancer treatment we would be endorsed.	Comments noted. The committee acknowledged that obtaining reliable clinical evidence for technologies for mental health such as depression can be challenging. It also noted that large inequities remain in treatments for mental health compared with other disease areas and considered this in its decision making. However, the costs and benefits of esketamine were very uncertain. Please see the updated guidance document (FAD) for further information which

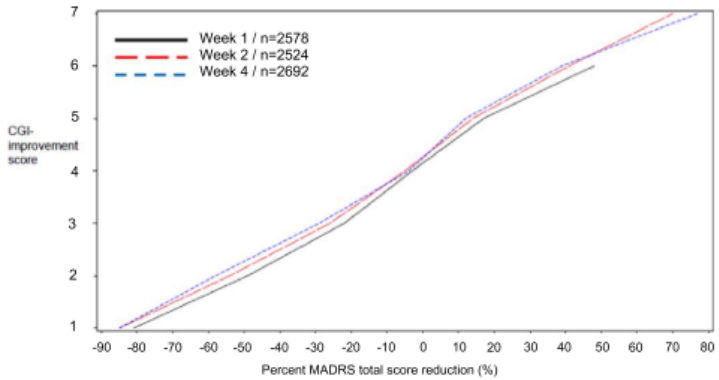
Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
				considered the issues submitted at consultation.
1	Commentator (web comments)	-	Has all of the relevant evidence been taken into account? Not sure Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? No - see comments Are the recommendations sound and a suitable basis for guidance to the NHS? No - see comments General Comments: We support the careful approach that you have taken to considering approval of this new treatment for depression, especially given its use as an anaesthetic agent and party drug with recognised abuse potential and association with considerable harm. We agree with the position of the committee that it does not seem to be simply a question of whether esketamine is cost-effective but whether it is effective at all. Depression is a long-term condition that can make people's lives harder. It therefore does not seem particularly relevant what the effects of any drug are after 4 weeks. This may be an appropriate time point for assessing response to an infection but not to mood states. There are a number of substances that might temporarily improve mood – like opioids, benzodiazepines, alcohol, cocaine and many other recreational substances- due to their effect of inducing a euphoric state. However, it is quite a different question as to whether these substances will produce a change that will be beneficial to a person in the long-term. We fear that short-term studies may demonstrate effects that are not borne out in the long-term – similar to that for many illicit substances, for which the long-term outcomes are generally dysphoric states. Indeed, this is the case for ketamine users who are general found to be dysphoric, even after they stop the drug (Morgan & Curran, 2012). It is therefore imperative that drugs which are provided in the NHS are rigorously tested for their long-term effects – both positive and negative. In addition to this, the effects of esketamine do not seem established even in the very short-term horizon of 4 weeks. Only one out of three short-term trials showed a significantly statistical effect on depression scores, and the effect was very small, with many people pointing out that it does not register as a clinically significant effect (C. Gastaldon, Papola, Ostuzzi, & Barbui, 2019; Horowitz & Moncrieff, 2020). We also understand that in two further as yet unpublished trials of esketamine in suicidal patients that the effect on depression scores was also not significant. It seems very unconvincing that this drug has any positive effects even in the short term. It seems rather concerning that although the company performed a 24 weeks study of placebo versus esketamine that they did not report the outcomes of this study in terms of depression scores. One wonders whether this is due to the results not aligning with their commercial objectives. Moreover, the effects of human contact seem to grossly outweigh the effects of esketamine in the trials. A 17-point reduction on the MADRS scale derived from salt water spray and time with a nurse seems to us a considerable effect, dwarfing the effect due to the drug (up to 4 points in some studies). We understand that it is highly unusual for a discontinuation study of a drug to be included in proving a drug works when the drug is	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			known to cause dependence and withdrawal as ketamine is. This study design would be deliberately confusing because withdrawal effects from esketamine would mimic symptoms of 'relapse' and so make coming off the drug look detrimental. We also understand that the company did not report withdrawal effects in its results meaning that it is not clear that this possibility was excluded. It would seem that the results from this study are not reliable as a result. Even not taking this into account, the difference between continuing or discontinuing do not seem to be very different at all after a few weeks suggesting that the drug is not really effective at preventing 'relapses.' Lastly the harms of this drug have been downplayed, but there is good reason to think that they could be quite significant. Ketamine is known to cause a number of health issues in recreational users or in patients given it in anaesthetic doses (admittedly larger doses than employed her, but notably patients will be given esketamine much more often than in anaesthetic practice). It is known to cause ketamine bladder, whereby the bladder wall is worn away over time, leading to people needing catheters to pass urine. It can cause heart attacks, and strokes due to the increase in blood pressure ('spikes'). It can cause motor vehicle accidents because ketamine has profound effects on hand-eye co-ordination, judgement and decision making. It has also been associated with suicides. This may be due to the psychotic symptoms(Beck et al., 2020; Wood et al., 2011) it is known to cause at sub-anaesthetic doses (including the doses tested by Janssen). It may be due to withdrawal effects from the drug (Schatzberg, 2019). It is hard to know the exact reason but it is surely very concerning that all these events occurred in the esketamine arm of Janssen's studies more frequently (and sometimes exclusively, in the example of suicides) in the esketamine arm compared with the placebo arm (C. Gastaldon et al., 2019; Horowitz & Moncrieff, 2020). It is perhaps more concerning to see these trends for harms extend into real-world practice with the same group of harms occurring in US in the year since the drug has been approved for use (Chiara Gastaldon & Kane, 2020). It is also alarming that the doses of esketamine used in clinical trials have also been shown to alter neurodevelopmental pathways in animal models leading to severe cognitive and behavioural impairments(Zimmermann, Richardson, & Baker, 2020). The long-term effect on the adult brain has not been investigated but these results are foreboding. This combination of factors – a lack of clear efficacy in the short-term, a lack of evidence of benefits in the long-term, very serious signals about harms from this drug, known risks of abuse and misuse – makes it seem reasonable to err on the side of caution and await more robust proof of efficacy, especially in the long-term (of a year or more) and verify that the safety signals are not likely to increase morbidity and mortality of users of esketamine (or other road users). It is also striking that the academics who have come out of support of this drug as having a 'novel action' or a 'breakthrough' are all paid by Janssen. For example, in a recent letter to the British Journal of Psychiatry (Kasper, Young, Vieta, Goodwin, & Meyer-Lindenberg, 2020), echoing the arguments put by Janssen to other critical papers, and echoing some of the submissions made to the NICE consultation process, all five authors receive money from Janssen. At least two authors are principal investigators on esketamine studies funded by Janssen. Their remarks about the 'novelty' of esketamine are reflected in submissions to NICE as outlined in the slides from the recent consultation hearing. The British Association of Psychopharmacology (BAP) has also made comments very supportive of the drug, without being supported by evidence. It has not been made clear whether the BAP has received direct payments from Janssen. Lastly, the clinical expert makes a number of points in the consultation document that are repeatedly favourable to Janssen's position (eg that a 4 point improvement on the MADRS represents a clinically significant difference, in contradiction of the existing evidence). There are likely to be few people with experience of ketamine in the UK given its limited deployment in clinical practice and many of the experts, such as Rupert McShane or Hamish McAllister-Williams (https://mood-disorders.co.uk/admin/resources/hamish-mcallister-williamsvns-and-restore-life17sept18.pdf), have close relationships with the manufacturer, including direct payments for consultancy as well as research support. We wonder whether the financial connections to the manufacturer of clinical experts involved in the committee's deliberations or that have been called on to give expert testimony might have influence the opinions presented, especially	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			when unsupported by the existing research evidence but consistent with the manufacturer's commercial objectives. Furthermore, it has been recognised that drug manufacturers often use the small Scottish market to put pressure on NICE to generate a favourable review in order to get access to the more lucrative English market. This seems to be occurring in this case with Janssen where they have agreed to subsidise esketamine in Scotland to render it 'cost-effective.' We hope that such political machinations will not influence the committee's appraisal of the scientific evidence. On the other hand analysis of esketamine by experts independent of financial ties to the manufacturer has unanimously concluded that the drug is not effective and its safety has been questioned (Cristea & Naudet, 2019; C. Gastaldon et al., 2019; Chiara Gastaldon & Kane, 2020; Horowitz & Moncrieff, 2020; Schatzberg, 2019; Turner, 2019). It is also notable that the national health evaluators in France, Denmark and Sweden have generally given a negative evaluation of esketamine's usefulness and not approved it for widespread use (although some countries have approved it as a fourth line antidepressants in some cases). Overall, given a lack of evidence for the effectiveness of this medication, a lack of long-term data and worrying danger signals, it would seem mandatory to demand greater evidence of safety and effectiveness before approving this potentially harmful treatment for widespread use. References Beck, K., Hindley, G., Borgan, F., Ginestet, C., McCutcheon, R., Brugger, S., … Howes, O. D. (2020). Association of Ketamine With Psychiatric Symptoms and Implications for Its Therapeutic Use and for Understanding Schizophrenia: A Systematic Review and Meta-analysis. JAMA Network Open, 3(5), e204693–e204693. https://doi.org/10.1001/jamanetworkopen.2020.4693 Cristea, I., & Naudet, F. (2019). US Food and Drug Administration approval of esketamine and brexanolone. The Lancet Psychiatry, 1–3. https://doi.org/10.1016/ S2215-0366(19)30292-5 Gastaldon, C., Papola, D., Ostuzzi, G., & Barbui, C. (2019). Esketamine for treatment resistant depression: A trick of smoke and mirrors. Epidemiology and Psychiatric Sciences, 08(Ci), 91–94. https://doi.org/10.1017/S2045796019000751 Gastaldon, Chiara, & Kane, J. M. (2020). Post-Marketing Safety Concerns with Esketamine : A Disproportionality Analysis of Spontaneous Reports Submitted to the FDA Adverse Event Reporting System. https://doi.org/10.1159/000510703 Horowitz, M. A., & Moncrieff, J. (2020). Are we repeating mistakes of the past? A review of the evidence for esketamine. The British Journal of Psychiatry, 1–4. https://doi.org/10.1192/bjp.2020.89 Kasper, S., Young, A., Vieta, E., Goodwin, G., & Meyer-Lindenberg, A. (2020). Letter to BJPsych in response to Horowitz and Moncrieff. British Journal of Psychiatry. https://doi.org/https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/are-we-repeating-mistakes-of-the-past-a-review-of-the-evidence-for-esketamine/C4BDC70050164FD9D88DF40967C2853A#fndtn-comments Morgan, C. J. A., & Curran, H. V. (2012). Ketamine use: A review. Addiction, 107(1), 27–38. https://doi.org/10.1111/j.1360-0443.2011.03576.x Schatzberg, A. F. (2019). A word to the wise about intranasal esketamine. American Journal of Psychiatry, 176(6), 422–424. https://doi.org/10.1176/appi.ajp.2019.19040423 Turner, E. H. (2019). Comment Esketamine for treatment-resistant depression : seven concerns about efficacy and FDA approval. The Lancet Psychiatry, 1–2. https://doi.org/10.1016/S2215-0366(19)30394-3 Wood, D., Cottrell, A., Baker, S. C., Southgate, J., Harris, M., Fulford, S., … Gillatt, D. (2011). Recreational ketamine: From pleasure to pain. BJU International, 107(12), 1881–1884. https://doi.org/10.1111/j.1464-410X.2010.10031.x Zimmermann, K. S., Richardson, R., & Baker, K. D. (2020). Esketamine as a treatment for paediatric depression: questions of safety and efficacy. The Lancet Psychiatry, 0366(19), 10–11. https://doi.org/10.1016/s2215-0366(19)30521-8	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
1	Commentator (web comments)	-	General Comment: As an ex Expert Clinical Assessor for MHRA, I fully support not approving esketamine. The reasons are laid out in a BMJ essay: The trouble with antidepressants: why the evidence overplays benefits and underplays risks--an essay by John B Warren http://bmj.com/cgi/content/full/bmj.m3200	Comments noted. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	Has all of the relevant evidence been taken into account? In clinical practice, it takes a substantial amount of time (measured in years rather than months) for a patient to trial (at a therapeutic dose) the currently available different types of oral anti-depressants before ECT might be considered. This leaves a gap in the available depression treatment care pathway. New treatments for treatment-resistance depression are therefore urgently needed to reduce the burden of depression for the patient and carers. Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? In my experience as Neuromodulation Lead Nurse for [redacted] and [redacted], the higher treatment burden claimed for Esketamine is no different (or possibly less) then the costs associated with ECT or Transcranial Magnetic Stimulation (TMS). Also, there are no significant cost issues (other than initial staff training) in using existing ECT clinics for Esketamine treatment. ECT clinics already have beds, monitoring equipment and staff in place. In addition, many other existing community mental health clinics (e.g. those set up for monitoring olanzapine depots) could be easily adapted to additionally monitor Esketamine treatment. Are the recommendations sound and a suitable basis for guidance to the NHS? The current conclusion does not currently take into account the inordinate length of time a patient must currently endure to receive appropriate treatment. If Esketamine can be proven to substantially reduce this time period then it should be considered for use in the NHS, especially as I believe that the costs estimates for Esketamine have been exaggerated by not including the considerable costs of ECT which is where such patients are often finally treated.	Comments noted. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	**Has all of the relevant evidence been taken into account?** See below **Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?** See below **Are the recommendations sound and a suitable basis for guidance to the NHS?** See below **General Comments:** Thank you for your thoughtful consideration of the issues relating to the approval of esketamine. We offer some further points of clarification regarding the Appraisal consultation document:	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the

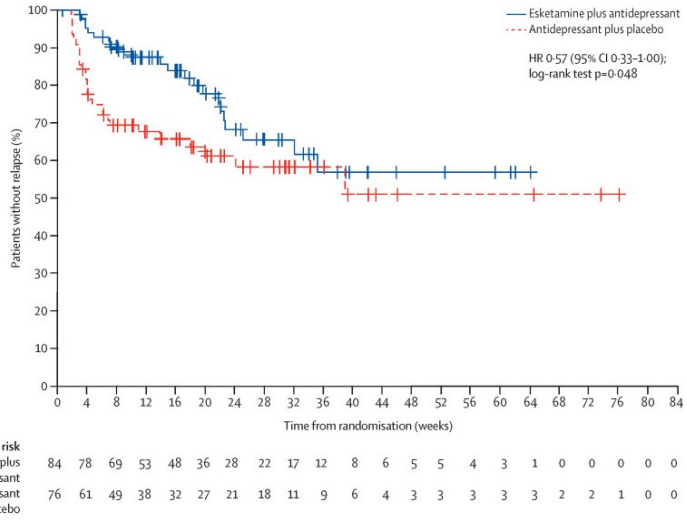
Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			3.1 Treatment-resistant depression. The characterisation of 'treatment-resistant depression' as severe or burdensome is not necessarily accurate. 'Treatment-resistance' refers more to the drug treatment of depression than the nature of the condition itself (and is borrowed inappropriately from the concept of 'antibiotic resistance'). Failing two antidepressants, which themselves have marginal efficacy,(Jakobsen et al., 2017) (as explained in the BMJ recently)(Warren, 2020) is not necessarily evidence of a severe condition. Furthermore, the entry criteria for the Janssen studies excluded people with suicidal thoughts, past history of ECT and co-morbidities so it is not clear that esketamine has been tested in a group that corresponds to most clinicians' idea of 'treatment-resistant depression' or 'severe depression' entails. Most widely used definitions of 'treatment-resistant depression' include patients who have failed numerous different antidepressants from different classes (often explicitly including MAOIs, TCAs or SNRIs), have psychotic features, have trialled ECT (many staging models of 'treatment-resistant depression' take into account the number of sessions of ECT received), have trialled antipsychotics, and psychotropic augmentation strategies (Ruhé, van Rooijen, Spijker, Peeters, & Schene, 2012). Therefore, the patients included in the trial would not fit most clinicians' impression of what severe or 'treatment-resistant depression' constitutes. This impression is further emphasised by the fact that this group of patients showed a very significant improvement to placebo (17-point MADRS reduction). Even if the TRD concept does represent something more severe, experience with numerous other drugs shows that the treatment is likely to be extended to many people with less severe conditions. 3.2 Unmet needs. Therapeutic hopelessness arises when pursuing ineffective treatments that have received hype from the marketing arms of their manufacturers, often echoed in views expressed by often-conflicted academics and similarly conflicted patient support groups (Fabbri, Lai, Grundy, & Bero, 2018; Fabbri et al., 2020; Lock, Seele, & Heath, 2016). Adding another ineffective medicine, associated with much manufactured hype will only contribute to hopelessness (see patient accounts below in Appendix). It is also a concern that patients have internalised 'treatment-resistance' as a self-descriptor, when it reflects treatment efficacy, rather than the patients themselves. 3.6 The effect of psychological therapy. The clinical expert has suggested that "the lack of inclusion of psychotherapy in the studies of esketamine appears to be setting a hurdle that not a single currently recommended pharmacological treatment has surpassed." This may be true regarding regulatory approval, but with regards practice guidelines such as those set out by NICE, this is untrue – other classes of antidepressants have been carefully compared to psychological therapy. As psychotherapy is an intervention that is as effective as most drug treatments in the short term, but which some evidence suggests is more effective in the long-term and maintains its effects after the end of treatment, and has fewer side effects than drug treatments it would be more prudent to raise the bar for the evaluation of drug treatments to require psychotherapy as a comparison, rather than lower the bar for esketamine. Although the clinical expert has suggested that some people have a condition which is too severe to be candidates for therapy, the fact that the group recruited for these studies showed a 17-point reduction in MADRS scores from passive social contact suggests that this would be a group that might respond to psychotherapy and that this comparison should be given consideration. 3.7 Efficacy of esketamine. The committee recognised that there were two negative trials regarding the efficacy of esketamine compared with placebo. We now	updated guidance document (FAD) for further information.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			know there are a further three trials that demonstrate no statistically significant difference between placebo and esketamine: ASPIRE-1(Fu et al., 2019) and ASPIRE-2(Ionescu et al., 2019), and Canuso et al (2018)(Canuso et al., 2018). The ASPIRE trials were conducted by Janssen in suicidal patients meeting a diagnosis of MDD, and have so far only been presented as posters as conferences. They both found no statistically significant difference between placebo and esketamine groups at day 25 and no difference in suicidality at 24 hours, questioning the claim that esketamine is effective for suicidal thoughts (Fu et al., 2019; Ionescu et al., 2019)). Overall, there are therefore five negative trials (on MADRS score at day 25), compared to one short term trial with statistically significant effects (TRANSFORM-2) and one discontinuation trial (SUSTAIN-1). 3.8 MADRS score We note the following: “The clinical expert noted that MADRS is non-linear, meaning that a change in score at the lower end of the scale does not mean the same, in terms of clinical importance, as a change in score at higher end of the scale.” We are not certain what this claim is based on, but analysis of MADRS and CGI scores demonstrates a linear relationship between MADRS scores and clinical impressions of overall improvement as seen in the figure below from Leucht et al (2017)(Leucht et al., 2017).  **Fig. 2.** Linking the percentage change from baseline in the MADRS total scores with the CGI-I scores. As the committee recognises, the definition of ‘remission’ and ‘response’, while widely used in academic studies, are by no means standardised or intuitively recognisable entities. Indeed, as mentioned previously the artificial dichotomisation of a continuous scale like the MADRS has the tendency to exaggerate differences between similar groups(Kirsch & Moncrieff, 2007), as appears to have happened in the case of esketamine. 3.9 Time period	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			Although the committee recognises that 4 weeks trials have little relevance to real world treatment of depression, they comment that further improvements may occur after the 4-week period. This is possible – however, it is more likely that effects will reverse after a period of treatment, as ingestion of psychoactive substances associated with tolerance and withdrawal often do – as in the case of opioids (which have been identified as having some overlap with the mode of action of esketamine)(Schatzberg, 2019) or benzodiazepines, both of which have diminishing effects over time due to tolerance. We note the BAP has submitted the comment "patients with TRD generally maintained their improvements seen at the end of acute treatment, and even on average improved further," But this statement is not supported by existing evidence. Indeed, the evidence from recreational ketamine users is that long-term use is associated with dysphoria. Increased depression scores were found in both daily users and ex-ketamine users over the course of 1 year(Morgan, Muetzelfeldt, & Curran, 2010), although not in current infrequent users (<3 times per week), so it is possible that this will not apply to esketamine users, but it also seems improbable to predict further improvement based on the common finding of dysphoria in long-term users of this drug (and many other similar substances). 3.10 Power/clinically significant difference. "The clinical expert commented that for a population in a trial, a mean difference of 4 was clinically significant," but there is no evidence for this statement. A change in score of 4 or less has been clearly shown to be less than the change required for a clinician to detect a 'minimally improved' difference, by the acknowledged leader in the field of establishing minimally clinically significant improvements (Leucht et al., 2017). In this study 22 drug trials, involving 3288 patients were analyses at several time points, with a highly consistent relationship found between MADRS score and CGI-I (an intuitive scale with high inter-rater reliability), finding that the minimally detectable difference by a clinician was 7-9 points. We are not aware of other analyses that have provided analysis on the clinical significance of change in MADRS scores, so it is not clear that it is accurate to say that "there is debate about what is considered a minimal clinically significant difference in the literature." Additionally, as the committee has recognised, the studies were powered to find a minimum difference of 6.5 MADRS point between the placebo and esketamine groups, and this means the study may be under-powered to detect a difference as small as 4 MADRS points, indicating the findings may not be reliable. The response to placebo was indeed large in the esketamine trials. This is perhaps unsurprising given that participants had several hours of contact per week (at least two hours with staff, two times a week, during administration of esketamine and supervision afterwards). Indeed, one interpretation of these trials is that human contact for several hours a week has a large effect on depressed people – a 17 point reduction in MADRS score corresponds to 'much improved' (Leucht et al., 2017). It does not seem reasonable however, to consider that a large placebo response to make it more difficult to ascertain the effects of esketamine: the central premise of a placebo-controlled trial is to identify drug-specific effects; there is no reason to suppose that this was not achieved in this trial. Regression to the mean is surely a factor in the trial – but again it would affect both arms of the trials equally. The EMA provided no evidence for considering the effect size to be clinically significant and it is unclear what their opinion was derived from.	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			3.11 Withdrawal design. The committee recognises that participants in SUSTAIN-1 were more likely to tolerate the adverse effects of the drug, but additionally they were also selected for inclusion only if they achieved treatment response (>50% reduction in baseline MADRS scores) in the short term trials.(Daly et al., 2018) Therefore the group selected for this trial represent an enriched group of patients who respond positively to the medication and do not represent the wider group of patients who would be given the drug in practice (ie not already selected for tolerability and response), who would be unlikely to show as large an effect. As a consequence of choosing 'responders', withdrawal designs have been described as tautological because they test whether a treatment shows an effect in a group who have been selected because the treatment shows an effect (that is, it is only 'responders' who are recruited into these trials).(Ghaemi & Selker, 2017) It is therefore likely that esketamine would have lesser effects in a less selected population. As regards unblinding, it is highly implausible, given the immediate physical and mental alterations produced by esketamine (a doubly potent enantiomer of a drug used to produce a 'high' for recreational users, at similar doses to that employed in the current trials) that any esketamine trial using a pharmacologically inert placebo could be truly blind. This also applies to withdrawal trials, where participants randomized to placebo will undoubtedly notice that they do not experience the same immediate effects as they experienced before. The FDA analysis of dissociation symptoms (using the CADSS) confirms this (p.28-29 of NDA)(FDA, 2019). The FDA found (in Figure 6 on p.28 of the NDA) that CADSS scores declined rapidly in the arm randomised to placebo. CADSS score was found to significantly associated with time to relapse of depression. The FDA offered several alternative interpretations, the more probable of which was that "the subject may worsen either due to suspecting they are no longer taking active drug." They also offer another explanation, which is also plausible: that the dissociative effects of esketamine are responsible for slight changes on depression scores because people are literally 'out of it', which may reduce the intensity of some domains recorded in the MADRS; when this effect abates, so does the marginal reduction on depression scores. 3.12 Withdrawal effects The point made by the committee in 3.11: "The committee also noted that people with depression in stable response or remission from the TRANSFORM trials who only had placebo had a lower relapse rate than those who stopped esketamine, although this was not explored fully by the company." The fact that relapse rates were higher in the group discontinued from esketamine than in the untreated group emphasises the fact that removing esketamine is not just revealing the underlying condition (which should produce relapse rates that are equal to the untreated group), but that there is an additional drug withdrawal effect. This suggests that many 'relapses' are in fact withdrawal effects, which have been mis-classified as 'relapse' because of the overlap between withdrawal effects and domains on the MADRS, as noted by the committee. This interpretation is strongly supported by the pattern of relapses which is visibly seen in the Kaplan-Meier plot reproduced below from a letter by the company in the Lancet Psychiatry(Singh et al., 2020). This figure is notable because it shows any separation between the two arms occurs in the first 4 weeks and that the curves actually cross after 36 weeks with barely any difference between the two. This pattern (large differences at short time periods, converging at long time periods) is the hallmark of a withdrawal syndrome: withdrawal symptoms cluster towards the point of cessation and slowly resolve over weeks or months. "The company considered that there would be no long-term withdrawal effects of esketamine because at this dose it leaves the body quickly." This is inaccurate. Withdrawal effects arise because the body has adapted (become tolerant) to the presence of the drug	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			(perhaps up-regulating NMDA receptors in the case of an NDMA antagonist, although there are many potential neurobiological mechanisms of tolerance/adaptation to esketamine) which is then removed. Withdrawal effects are the subjective experience of a system which has become homeostatically 'tuned' to a certain level of drug then being subjected to lower levels of it. The more quickly a drug leaves the body the more quickly and severe will be the withdrawal effects. For example, short acting benzodiazepines, antidepressants with short half-lives (eg paroxetine) and opioids with short half-lives produce the most severe withdrawal symptoms. *Withdrawal symptoms persist for the period of time taken for adaptations in the brain and body to resolve back to a pre-drug state – not for the period of time taken for a drug to leave the body.* This process of resolving adaptation can take weeks or months (although possibly less after short-term use). It is therefore misleading to suggest that a short half-life for a drug is consistent with there being no withdrawal effects; it is the opposite that is true. The clinical expert is misleading when he states: "that withdrawal effects of ketamine seen in recreational use are from higher doses." In fact, doses of esketamine employed in the trials are similar to those used recreationally. Esketamine doses employed by Janssen (56-84mg), were not distinct from those used by recreational users (equivalent to 50-100mg esketamine),(Sassano-Higgins, Baron, Juarez, Esmaili, & Gold, 2016) noting that esketamine has twice the potency of ketamine. Moreover, we know from experience with other drugs that, there is no threshold for which withdrawal symptoms occur and no reason to think they would not occur at the doses of esketamine employed by Janssen. Although 'sweating and shaking' may be more common with withdrawal from higher doses, withdrawal commonly involves numerous other symptoms such as fatigue, poor appetite, drowsiness, anxiety and dysphoria, which, as the committee note, overlap with items on the MADRS and may be mistaken for a relapse of depression.(Chen, Huang, & Lin, 2014; Cosci & Chouinard, 2020; Morgan & Curran, 2012)	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			 **Figure Kaplan-Meier estimates of time to relapse among patients who achieved stable remission** Hide caption 3.13 Outlier site. We maintain that it is reasonable to exclude the outlier site and we point out that the EMA did not provide any rationale for retaining it. The data is highly inconsistent with the data from other sites. Analysis of this study without this site showed there was no significant difference between the two arms, using Fisher's exact test (p=0.13).(Turner, 2020) This test is more appropriate than the log-rank method preferred by the drug manufacturer, because the log-rank method exaggerates the effect of early relapses, inappropriate in this case where these relapses are most likely to reflect withdrawal-related events.(Singh et al., 2020) The sponsor's re-analysis, using the log-rank method, excluding the outlier site yielded a barely significant P=0.048, but one wonders whether the FDA's analysis maintained significance, as they stated several times in the regulatory review documents that the outlier site "drives" the overall study results.(Turner, 2020) Although patient-level data was requested in order to resolve this issue, the manufacturer has not, to our knowledge, furnished independent researchers access to this. 3.14 Exclusions. The manufacturer excluded patients who would meet the usual clinical definition of 'severe depression', including suicidal ideation (in the last 6 months), suicide attempts (in the last 12 months), co-morbidity (drug and alcohol and mental conditions) and past history of ECT. It would seem that the patient population was a much less severe group of patients than clinicians would generally consider to be 'severely depressed' or 'treatment resistant'. This notion is emphasised by the fact that these 'treatment resistant' patients showed a large	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			response (17 MADRS points) to placebo (human contact). This makes the lack of clinical efficacy and the numerous safety signals for esketamine even more concerning – and the fact that a number of these patients committed suicide, despite having been selected for lack of suicidality. 3.15 Placebo adjustment The suggestion by the manufacturer that the placebo arm of a placebo-controlled trials is adjusted down violates the basic premise of placebo-controlled trials. Every point raised by the company as to why placebo response rates were high also relate to the esketamine arm. 3.16 Safety A salient point was made in the BMJ this month by John Warren, former Expert Medical Assessor and NDA evaluator for the MHRA: that in trials, manufacturers use composite scores in order to find positive effects for their drugs but use separate events for side effects, which has the effect of minimising the overall burden of negative effects.(Warren, 2020) This was evident in the esketamine trials, which was identified as a case in point. MADRS is a composite score which measures 10 symptom domains, including appetite changes, mood, sleep etc. Composite measures are more likely to find differences between groups (as small effects add up). In contrast, the many 'side effects' of esketamine were grouped individually. As stated in the BMJ article: "An incidence of at least 5% and at least twice that of placebo was reported for dissociation, dizziness, nausea, sedation, vertigo, hypoaesthesia, anxiety, lethargy, blood pressure increase, vomiting, and feeling drunk.(8) Whereas one primary endpoint was used to summarise benefit, safety was analysed as a collection of symptoms with no single endpoint, mitigating against finding statistical significance (32) and leading to the asymmetrical analysis of risk and benefit.(33)" Post-marketing surveillance of esketamine has only served to strengthen these concerns with strong signals emerging for the data for an increased risk of dissociation, sedation, feeling drunk, suicidal ideation and completed suicide(Gastaldon, Raschi, Kane, Barbui, & Schoretsanitis, 2020). Although such post-marketing surveillance data are inherently limited in drawing conclusions about causality, because they are not randomised participants, confounding by indication is possible, and there may be a notoriety bias in reporting, many of these effects (including suicidal ideation) remained when comparisons were made to adverse effects reported for venlafaxine and, furthermore, the effects reported closely mirror those reported in the regulatory trials submitted to the FDA, triangulating the evidence. The increased rates of worsening of depression and suicidal ideation in the esketamine group compared with the placebo group in the regulatory trials, although small in number, is another signal consistent with the risk for deterioration and increased suicidality in esketamine use. This may well be explained by the intense dysphoria that some users can experience during and after short-term treatment (see patient extract below in Appendices) and which is reported to occur with long-term use among recreational users (Morgan & Curran, 2012) One important issue not as remarked upon in the otherwise comprehensive appraisal is the strong association of motor vehicle accidents with esketamine use. Esketamine is known to impair hand-eye coordination, to cause dissociation, and to be associated with car accidents(Cheng, Chan, & Mok, 2005; Morgan & Curran, 2012; Schifano, Corkery, Oyefeso, Tonia, & Ghodse, 2008). In the regulatory trials there were 5 car accidents all in the esketamine arm, one of which was fatal. As can be seen in the supplementary material of	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			Gastaldon et al. (2020) a number of road traffic accidents were attributed to esketamine in the FDA database. In Hong Kong where recreational ketamine achieved particular popularity in the 1990s 9% of all fatal traffic accidents from 1996 to 2000 involved ketamine use (Cheng et al., 2005). There are also some elements of the safety study design (SUSTAIN-2) (Wajs et al., 2020)that do not lend it to a full understanding of the adverse effects of esketamine. 802 patients were enrolled into this study but 331 were excluded as the 'study was stopped by sponsor'. The explanation given for excluding almost half the patients was unclear. In the remaining patients there were some concerning findings: 114 patients had new onset suicidal ideation (in a group selected for lacking any suicidality), there were 6 suicide attempts and one completed suicide (in a patient with no history of suicidal behaviour or intent). There is some speculation that the increased suicidality seen in esketamine use may be related to its psychedelic properties (which may be useful for some, but lead to terror and fear for others). The dissociation caused by esketamine may be one manner of describing the hallucinogenic and psychedelic properties of the drugs. Lastly, there is concern that NMDA antagonists can be neurotoxic in the long term. Ketamine was originally developed from phenylcyclidine (PCP), known as 'angel dust' when used recreationally. Ketamine and PCP have similar chemical structures, and are both primarily NMDA antagonists although the potency of PCP is greater than ketamine. PCP causes similar effects to ketamine: hallucinations, distorted perception of sound (see patient accounts below). It also causes an increased risk of suicide, which some have linked to its ability to produce flashbacks (and some have linked ketamine suicides to flashbacks as well). It is also used as an anaesthetic agent and can cause euphoria in the short-term. It is, like ketamine, psychotomimetic (Murrie, Lappin, Large, & Sara, 2020). Both PCP and ketamine are highly lipid soluble, and ketamine possesses a chlorine atom (halogenation is widely used in anaesthetics to enhance penetration of the brain as it dramatically increases transfer of the molecule across cell membranes). The dangers of NMDA antagonists like ketamine have been demonstrated that repeated exposure to ketamine-like drugs during development can permanently disrupt neurodevelopment and have catastrophic long-term cognitive and behavioural outcomes (Zimmermann, Richardson, & Baker, 2020). In animal models, exposure to NMDA receptor antagonists during development in animal models impairs parvalbumin maturation, reduces the number of parvalbumin neurons in the medial prefrontal cortex and causes disorganised prefrontal cortex output in adulthood, mimicking the disease pathology of schizophrenia (Zimmermann et al., 2020). Notably, the doses used to induce schizophrenia-like dissociative symptoms and disrupt parvalbumin development in animals are similar to the doses used to in the Janssen depression trials (Zimmermann et al., 2020).This is reflected in its pro-psychotic effects in adult subjects given doses of esketamine used in the depression trials (Beck et al., 2020), although possibly the neurotoxic effects of esketamine might be less potent in adult brains. However, this possibility has not been excluded with long-term safety trials focusing particularly on brain effects. The cognitive impairment and dysphoria seen in recreational ketamine users is not reassuring (Morgan & Curran, 2012). Placebo-controlled long-term studies of safety utilising composite scales to assess for a range of safety side effects, with particular focus on neural effects of long-term administration of esketamine are required for the medication is released into general use. A Ketamine Side Effect Tool and Ketamine Safety Screening Tool have been developed by ketamine experts in Australia for this purpose(Short, Fong, Galvez, Shelker, & Loo, 2018). 13 major side effects are included based on a systematic review of the literature, including headache, dizziness, dissociation, increased blood pressures, blurred vision, nausea, sedation or drowsiness, faintness or light-headedness, anxiety, elevated heart rate, cognition side effects, urinary tract side effects and dependency risk (Short et al., 2018). Although severity ratings would be required to adequately match the level of detail captured in composite measures of efficacy such as the MADRS, an	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			initial assessment of data collected by Janssen using this composite scale would be instructive to compare with the positive effects of the drug. 3.22 Adjustment to mortality It was also concerning to see the manufacturer attempt to suggest that mortality would be improved, when there were more deaths in the esketamine group than the placebo group (even taking into account that patients spent longer on esketamine than placebo). There was barely any discernible effect of esketamine on depression scores to justify the calculation they proposed, and projection over time of results found in 4 week trials is unreasonable especially given the experience in recreational use with ketamine that it is associated with dysphoria in the long-term (Morgan & Curran, 2012). 3.25 Stopping treatment As with all drugs that cause dependence and withdrawal, and can cause addiction, it will likely be hard for people to stop esketamine after a period of use, which will contribute to long-term use. Some of these patients might resort to buying ketamine off the streets as occurred with opioids and heroin in the US – despite the view offered by the manufacturer that addiction is not possible because use is supervised. Addicted people will find a means to obtain a supply. As has happened with benzodiazepines, and now antidepressants, people who stop their esketamine are likely to experience withdrawal symptoms and as the manufacturer is informing doctors and patients that withdrawal effects are not possible, patients will be diagnosed with 'relapse' of their condition, and will be told to re-start treatment. In this manner many people will be caught in long-term use by withdrawal symptoms – a very common phenomenon in psychiatry as highlighted by the recent PHE report on prescribed drug dependence(Public Health England, 2019) in which they said: "Recurring patterns are evident in the history of medicines that may cause dependence or withdrawal. New medicines are seen as an important part of the solution to a condition, resulting in widespread use. Their dependence or withdrawal potential are either unknown at this point, due to a lack of research, or perhaps downplayed. As evidence of harm from dependence or withdrawal emerges, efforts are made to curtail prescribing. The repetition of this pattern is striking." 3.33 Innovative action 'Innovative action' is not a positive thing in its own right in the absence of evidence for meaningful efficacy and safety. This is a marketing device, not a serious point. It is like saying 'we have installed new breaks on your car: they are a totally new design! But they don't work very well and sometimes catch on fire.' In relation to the point "esketamine is sprayed in the nose which means it works rapidly": depression is a chronic condition and the ability to quickly absorb something through the nose is of no consequence in the time scale that matters. Other anaesthetic agents such as propofol (Mickey et al., 2018) and nitrous oxide (Nagele et al., 2015) also reduce depression scores in a few hours, but it seems improbable that this represents a sustainable effect on a long-term condition and is likely broadly similar to the effects of esketamine.	
1	Commentator (web	-	Has all of the relevant evidence been taken into account? I'm afraid I do not think so. As I don't think the evidence related to the direct and indirect cost of depression has been fully appreciated by	Comments noted. Following a

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
	comments)		the committee. Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? No. There is a huge unmet need for patients who fail to respond to traditional antidepressants. According to the available evidence I believe that Esketamine is cost-effective if prescribed to the right group of patients. Are the recommendations sound and a suitable basis for guidance to the NHS? No	second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	Has all of the relevant evidence been taken into account? Yes General comment As a consultant psychiatrist I write to recommend that this medicine be given approval in the England and Wales and Northern Ireland, following the example set by the Scottish Medicines Consortium earlier this year. I have worked for many years with patients suffering from therapy resistant depression, which is a totally debilitating condition, meaning that otherwise intelligent people are required to be signed off from their workplaces for years at a time. One of our patients has been part of Phase 3 clinical trial on Esketamine nasal spray. The effect of this trial has been life changing and has completely reversed the disorder, and our patient is now able to benefit from the talking therapy she has been undertaking throughout her disorder, as she can now feel the benefits of the work she has been doing on her mood, rather than just intellectually understanding what her therapist has been discussing with her, but it having no effect on her mood. I respectfully submit to the committee that this medication should be approved for use in England Wales and Northern Ireland for the benefit of those receiving treatment on the NHS.	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	General comment I feel that with Covid being a likely ongoing issue, that this may be an alternative to ECT. ECT could not be offered in a Covid safe way, but this may offer a realistic alternative using the same staffing and resources.	Comments noted. Following a second round of consultation

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
				on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	**Has all of the relevant evidence been taken into account?** p13. 3.10 It seems unfair to criticise the trial design when the design was dictated for licensing by the MHRA and FDA. The FDA/MHRA did not allow a placebo group because TRD is such a serious condition that it would be considered unethical to allocate some people to a placebo. Treatment-Resistant Depression is a condition people live with every day and is so severe, horrible and painful that some find suicide the only option. https://www.theguardian.com/society/2020/sep/01/male-suicide-rate-england-wales-covid-19 There are very few other conditions where the symptoms are so unpleasant people commit suicide because of them. And many who do not end their lives only do so because of the effect that would have on their families. As you acknowledge, hopelessness is a predominant symptom of TRD and, quite often, it is a realistic and comprehendible hopelessness. page 13 3.11: trial design: We appreciate that it is right that the ERG will be concerned about adverse effects. However someone with TRD may be willing to tolerate a much higher side effect burden because the symptoms of TRD are so deeply unpleasant. Moderate side effects are a long way from the actual symptoms of TRD so please don't assume side effects will put people off. When you're in a hopeless worthless state transient side effects (as opposed to continued medicine taking) can be tolerated for the final relief. page 9 3.6 : We do not know where the ERG has got the idea that it is a problem that someone couldn't have CBT immediately after esketamine. We are not aware that anyone suggested it could be. ECT would have exactly the same "restriction". A distinct advantage of esketamine is that, while it produces side effects at the time, these are limited to the day of treatment which then leaves the person side effect-free on the other 5 or 6 days. This contrasts with oral antidepressants where side effects are continuous and may impede CBT's effectiveness. We would be fascinated to read a NICE appraisal of CBT with the same degree of forensic analysis of studies as it would be unlikely to prove CBT is cost effective. We feel it unlikely patients would even want CBT on a day they have had other treatments.	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			Trends Psychiatry Psychother 2020;42(1):92-101. Cognitive-behavioral therapy for treatment-resistant depression in adults and adolescents: a systematic review Stephanie Zakhour 1 , Antonio E Nardi 1 , Michelle Levitan 1 , Jose Carlos Appolinario 1 2 3 4 PMID: 32130308 **Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?** Page 3: See comment in section 1 regarding the FDA and MHRA licensing requirements. The analysis should identify that the licensing trials required by the FDA and MHRA led to them awarding a licence and that they did not see this as compromising proof of efficacy. It is true that people may be more unwell than in the clinical trials. This is true of almost every condition with a new therapy and TRD is no different. We are unclear on the significance of this statement Alternatives: Psychological therapies seem to get a disproportionate degree of prominence in the TA. It is worth knowing that there are few trials of CBT in TRD and it is inappropriate to mention CBT as if to suggest it were a comparable therapy. The most recent systematic review (Zakhour et al, 2020) identified 8 studies of CBT in TRD, although one was an open trial and one was a case report. The remaining six were Randomised Controlled Trials but used waiting list controls (as opposed to double-blind placebo-controlled studies; waiting list controls are where the intervention is compared with someone having no treatment at all, told they were ill and could be treated but have to go on a 16 week waiting list). You cannot account for the impact of being on the waiting list group (which NICE itself pointed out was worse than a placebo). These CBT trial designs are something that, as general medicine clinicians, you would not find acceptable as proof of efficacy in any general medical drug. Thus, although CBT is an option and clearly helps some people, it has little robustly proven efficacy in TRD, or none if you do not accept waiting lists as an adequate control group. page 3 point 1: there is no evidence that we are aware of that supports this statement. It would seem logically virtually inconceivable that improving TRD will NOT improve someone's quality of life. We feel that is a distracting uncertainty. The ERG even acknowledges the wider effects of TRD p5, 3.1: "The committee concluded that the condition has a negative effect on people, their families and carers." If TRD has that impact how could going into remission not improve QoL? We would request the ERG re-considers this statement, the evidence supporting it and that this distracts the reader. page 18 3.16: We think that the ERG has greatly over-emphasised the risks of diversion and abuse. Esketamine will, for many years, be managed by MH Trusts. It is a schedule 2 Controlled Drug, with records and stock control of purchase, supply, administration and destruction. The risk of diversion of a medicine in a countable pack (as opposed to e.g. a liquid in a stock bottle) is close to zero and would be rapidly identified even if only one single pack disappeared. To imply that MH Trusts could not manage this is incorrect. For someone to abuse esketamine you would first have to prove you had TRD to a Consultant Psychiatrist. You would have a first dose, be monitored, the dose stays stable and then reduces (unlike ketamine, where the dose is increased in order to reach the same effect), you can't take the medicine away with you, it is given under direct supervision, and ketamine is readily available on the streets. On discussion with a DrugScience colleague they felt abuse of esketamine under the strict management conditions highlighted above would be very unlikely with ketamine readily available at £25-30 per gram on the streets. page 6 3.3: It would be almost inconceivable that someone with TRD would have not had at least two of these antidepressants so the	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			option list looks rather more impressive than it is. They all have similar modes of action on serotonin and/or noradrenaline, and STAR*D shows that response decreases significantly as the number of antidepressants tried increases. page 11 3.9: We are surprised at how negatively this has been interpreted. A separation from placebo after 2 days in TRD? why is this not considered a true effect? A 21-point decrease in MADRS scores cf. baseline is, to us , not open to uncertainty. page 11-12, 3.9: we would challenge the 4 week comment in relation to patient experience. Going into remission within 4 weeks would be incredibly helpful to many patients with TRD. page 14, 3.12: withdrawal effects: We are pleased that the ERG accepted this. It is important that ketamine is seen as different in this respect, due to escalating doses used by people abusing ketamine and reducing doses for esketamine in depression (after the first dose). It is worth recalling that esketamine is already available on the UK market as an anaesthetic. P14-15 3.12: withdrawal effects would be difficult to distinguish. we do not support this statement and would be surprised if this was agreed with by those with mental health experience. clinically you would not expect to see long withdrawal symptoms lasting days and looking like depression from something with a relatively short half-life where the central effect is virtually gone from the body in a few hours and which is only taken once or twice a week, not continuously. Any such withdrawal symptoms would not in practice be confused with a relapse of depression. Withdrawal symptoms would change and reduce over the time between doses. Depression would not improve and would most likely get worse. page 18 3.15: We do not know why the ERG has got this idea that it is a problem that someone couldn't have CBT immediately after esketamine. We are not aware that anyone suggested it could be. ECT would have exactly the same "restriction". A distinct advantage of esketamine is that, while it produces side effects at the time, these are limited to the day of treatment (and essentially only for a couple of hours), which then leaves the person side effect-free on the other 5 or 6 days and more able to make the most of CBT. This contrasts with oral antidepressants where side effects are continuous and may impede CBT's effectiveness. P18 3.16: the issue of abuse and diversion again: There may be an increased risk but, as predominantly secondary care pharmacists, we do not recognise this as a risk that cannot be managed. Widespread diversion is hard to imagine and there are plenty of other drugs with higher abuse potential and higher availability. Schedule 2 drugs are widely used e.g. ketamine, diamorphine etc and whatever risk there is can be managed within existing systems. Use of standard Controlled Drug systems (especially with a single-use mechanism in distinct boxes) eliminate this risk. As a schedule 2 Controlled Drug, there are records and stock control of purchase, supply, administration and destruction. as said earlier the risk of diversion of a medicine in a countable pack (as opposed to e.g. a liquid in a stock bottle) is close to zero and would be rapidly identified even if only one single pack disappeared. page 30 3.17: We fully recognise that depression can be episodic and agree with the clinical expert but our aim is always remission and conventional antidepressants have a significant relapse prevention effect ("The average rate of relapse on placebo was 41% compared with 18% on active treatment; Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Geddes et al, Lancet 2003;361:653-61). If treated well it does not need to be episodic.	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			page 21 3.19: This would best be phrased “Treatment-resistant depression can be an episodic condition if inadequately treated. page 22 3.20: We would suggest a more accurate phrasing would be “treatment again if symptoms returned”. I think you’ll find this is true of any treatment and clinical area. P31 3.30: That may be true but experience from abroad shows that the dissociative effects can be significant for an hour or so but that, after the first treatment, wear off with subsequent treatments and rarely need anything more than reassurance plus a low or managed stimulus environment. The ERG should not underestimate or overestimate this and it will require some training but not a huge amount. training is required in general to understand and managed TRD better. page 31 3.30: We are concerned at the financial calculations in relation to the service impact. Secondary care mental health NHS Trusts are capable of safely managing CDs e.g. methylphenidate, methadone, buprenorphine, and opiates (all more abusable than esketamine) so the costs of managing a registry would be possible though we accept service pressures would need discussing and funding. Costings should also include savings from lower use of ECT which includes employing consultant anaesthetists and larger numbers of staff. P32 3.31: use of an ECT suite: We are unsure what this means. Trusts will not necessarily need to use ECT suites but could find a quiet area with appropriate monitoring and more homely and calm than the probable clinical sterility of an ECT suite. P33. 3.31: As highlighted in last response if 82% of MH Trusts have plans on how to implement esketamine this is a large proportion and gives an illustration of the need for an effective treatment for TRD. we agree this will be challenging however this is often the case in mental health services and feel not a reason to refuse a licensed treatment. Yes, it might take a bit longer but, as COVID-19 has shown, Trusts can implement changes extremely quickly when they want to. That only 18% did not have plans might now be out-of-date and is a low number. therefore an extended implementation would be needed for some STPs P36 3.34: This is true. Methadone, methylphenidate and buprenorphine are examples. P6 3.3: ECT personal comment: ECT is an option but is not a passive treatment. A colleague at Minds entire job at one time was supporting people pre-ECT and post-ECT with their fears, apprehension, distress, and memory loss. ECT also has a number of important contraindications and cautions: ECT: Contraindications Definition Before discussing contraindications, it is important to first understand the physiologic effects of ECT. These include: • Large increases in cerebral blood flow and intracranial pressure • Initial parasympathetic discharge manifested by bradycardia, occasional asystole, premature atrial and ventricular contraction, hypotension and salivation • Following parasympathetic reaction is a sympathetic discharge associated with tachycardia, hypertension, premature ventricular contractions, and rarely, ventricular tachycardia and ECG changes, including ST-segment depression and T-wave inversion, may also be seen. • Glucose homeostasis is also affected. Hyperglycaemia seen in insulin dependent patients	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			Absolute contraindications: • Known pheochromocytoma Relative contraindications: The risk of the patient's psychiatric illness, side effects of antidepressant medications must be weighed against the risk of ECT and anaesthesia. These conditions include: • Increased intracranial pressure, ok if there is not a mass effect • Brain tumours, same recommendation as above • Recent stroke- ECT has been performed successfully • Cardiovascular conduction defects. Pacemaker is not a contraindication to ECT- AICD function can be deactivated and magnet should be available if needed • High-risk pregnancy- OB consult and fetal monitoring is recommended • Aortic and cerebral aneurysms • Asthma/COPD- some suggest that you should discontinue theophylline because of its potential to cause status epileptics. Recommendations: • Delay ECT for patients with unstable angina, decompensated heart failure, or severe symptomatic valvular disease until these conditions are stabilized or optimized. Cardiology consultation may be of benefit • For high-risk neurosurgical lesions including recent stroke and brain tumour, neurosurgical consultation is recommended • Diabetic patients should hold oral hypoglycaemic, short acting insulin and halve their long acting dose with fasting • Warfarin can be continued in high risk patients with INR <3.5 • In severe GERD antacids can be taken or intubation considered. (Source https://www.openanesthesia.org/ect_contraindications/ accessed 10.9.20). **Are the recommendations sound and a suitable basis for guidance to the NHS?** We feel that this TA could read as if the chair and panel has decided to not recommend esketamine regardless of any evidence. Concentrating on uncertainty is a technique regularly used in politics and business, and was used by e.g. the tobacco industry to try to minimise the effects of overwhelming data (but not 100.00% certain) that smoking caused lung cancer. This emphasising of "doubt" was effective and allowed the tobacco industry to keep up sales (and deaths) for many decades. A similar campaign is being waged on climate change because no one can be 100.00% sure it's human caused and, if it isn't, we can carry on as we are. We think you are asking too much for a novel and innovative therapy for a life-threatening condition in the early stages of introduction. It is easy to come up with uncertainties but we like the committee to attempt to see through these and see the overall message. we would accept that esketamine could be 4th or 5th line in TRD (which is what will likely happen anyway) or that you give esketamine a limited approval with a further review in 2 years' time when more is known about the longer-term outcomes and more studies are published, and the registries show up, but would ask it is not turned down entirely because of the inevitable uncertainties about the finer details of the studies. Studies in such heterogenous conditions do not always come up with blanket treatment effects as you might for a more homogenous illness. Esketamine may be an on-going treatment for some, involving a day, a visit to a clinic (possibly many miles away, especially in the many rural areas), the need for an accompanying person or taxis, a treatment that is rather more than just popping a pill, and a significant routine. Not dissimilar to oncology treatments in fact and that can be for life-threatening conditions too. We would thus appeal to the ERG to just take a step back and think "we have a condition here that is so horrible people commit suicide, we have a treatment that clearly helps a significant number (even if there are some uncertainties about the deeper details), and the alternatives are lacking or crude and with low patient acceptability. Can we really deny severely depressed people a treatment that works	

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			when other treatments do not?" P24 3.22: Personal comment: A pharmacist in the UK who after 11 years TRD had decided to end it all as she could no longer bear the daily depressive symptoms but, after 3 treatments of esketamine (her last roll of the dice), phoned her husband to tell him where her secret stash of poisons was kept and told him to destroy them as she no longer wanted to end her life.	
1	Commentator (web comments)	-	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? There is adequate real- world evidence in the USA/ Canada where this useful treatment has been available for almost 2 years. We have clinical evidence from some centres such as Oxford and Northampton in the UK. I have listened to many podcast by other experts highlighting its benefits in treatment resistant depression. At the moment with all of our combination drug treatments there is a significant "time lag". Only ECT can "jump start" recovery but relapse rates are high and some patients need maintenance ECT for a long time. I believe that this "rapid" acting antidepressant with "high remission" rates at 4 weeks is similar to ECT at 4 weeks of biweekly treatment. This is far higher than in STARD level 3 (14%) and or 4 (13%). Our patients and we as clinicians should not be deprived of this novel break through treatment option. I am writing as an advocate of my patients who are not very vocal. I am hoping that the patient groups and professional organisations will make similar points. SMC has already approved its use in Scotland within its marketing authorisation. Why should English TRD patients be deprived of this excellent treatment option as an alternative or an intermediate step before ECT.?? If cost is an issue acute and long term then at lease restrict its use but not deny it to our vulnerable patients as a useful treatment option. Where has the patient choice gone? This is a major public health disorder with serious disability and personal consequences. In the health economic modelling you will need to take into account expensive and lengthy hospital stays for ECT and or other combination treatments. There must be some cost savings as this novel treatment can be administered on an out patient basis with home treatment support. Are the recommendations sound and a suitable basis for guidance to the NHS? Certainly not. In my personal opinion it should be available after 2 AD's have failed at adequate doses and adequate duration with good compliance. Depending on the severity of the illness. Very severely ill patients who might be considered for ECT must have this is an alternative. Less severely ill patients can go down the switch to Vortioxetine (within license) or augmentation treatments with Atypicals and or Lithium to start with. Esketamine could become a 4th line option before ECT.	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders. Please see the updated guidance document (FAD) for further information.
1	Commentator (web comments)	-	General Comments As a Consultant Psychiatrist in the NHS, frequently treating patients with treatment resistant depression (TRD), I am extremely disappointed with your decision to not recommend Esketamine with a serotonin selective reuptake inhibitor (SSRI) for this patient group. The most comprehensive study of MDD treatment resistance was the National Institute of Mental Health-funded Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. (ref 1) . In this study, patients with MDD underwent a series of sequential antidepressant treatments in monotherapy or combination, or psychotherapy trials using evidence based antidepressant treatment strategies. Acute remission rates decreased with each STAR*D level (level one 37%, level two 31%, level three 14% and level four 13%). Resistance to treatment becomes markedly increased at level 3 (after failure of two treatments), and predicts a poor prognosis with respect to future treatment efficacy and tolerance. Further more, patients with TRD experience relapse at a higher rate than do those with treatment-responsive MDD. Even when patients with TRD respond to treatment, the overall relapse rate while continuing treatment	Comments noted. Following a second round of consultation on draft guidance, the appraisal considered all submissions and comments from stakeholders.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment	NICE Response
			with the same antidepressant is high after 2 (65%; within 3.1 months) and 3 failed trials (71.1%; within 3.3 months).(Ref 1) . Therefore, there is a substantial unmet need for effective treatments that can sustain antidepressant benefits for this population with)TRD . New treatments with good evidence in short-term and in maintenance trials , such as Esketamine with a selective reuptake inhibitor (SSRI), are very much needed in clinical practice today in order to bring hope to our patients and alleviate their suffering. I very much hope this decision is changed. Conflict of Interest/Disclosures: I receive only a salary from a full-time NHS post. I have no shares or positions in the pharmaceutical industry. I have received in the past consultancy fees from most of the pharmaceutical companies based in the UK.	Please see the updated guidance document (FAD) for further information.

Esketamine for treatment resistant depression [ID1414]

Janssen response to NICE ACD 2 – 16th October 2020

Executive Summary

Janssen welcomes and thanks NICE for the opportunity to comment on the second Appraisal Consultation Document (ACD 2) for esketamine for treatment resistant depression [ID1414].

Overall, Janssen is disappointed with the decision not to recommend esketamine nasal spray (ESK-NS) for treatment-resistant depression (TRD) in ACD 2, despite the Committee's recognition of the unmet need for patients with TRD and the impact this condition has on patients, families and carers. As noted by the NICE Committee and a number of consultees, TRD is a seriously debilitating, potentially life-threatening condition. With each treatment failure, TRD disease morbidity increases, with reduced quality of life, increased costs and poorer outcomes observed (1). Unipolar major depression is projected to be the leading cause of disease burden by the year 2030 worldwide (2), with much of the overall burden of major depression falling on those with TRD (3). Newly licensed treatments which enable the optimal treatment of TRD are urgently needed given the limited innovation in the disease area for the last 30 years. Mental health represents up to 23% of the total burden of ill health in the UK but only 11% of NHS England's budget (4). Despite the Health and Social Care Act 2012 calling for 'parity of esteem' between mental health and physical health, large inequities remain. This is especially pertinent given the unprecedented COVID-19 pandemic, with increased levels of mental health support and investment required. Janssen therefore remains fully committed to addressing the Committee's concerns raised in ACD 2 and ensuring that patients with TRD and healthcare professionals have the option to access ESK-NS for TRD. This is especially important given the lack of innovative treatment options and inequities that persist in mental health compared to physical health conditions.

We note that the Committee's concerns in the ACD fall into 2 key areas: questions around the clinical effectiveness of ESK-NS and the economic model. We understand that some of the clinical concerns have informed the Committee's conclusions regarding the assumptions made in the economic model. Our response therefore addresses the clinical concerns raised by the Committee first, before addressing the economic modelling issues, and is summarised below.

In addition, we are aware of the Committee's points in the ACD around the positioning of ESK-NS and some of the Committee's uncertainty regarding the clinical and economic evidence. We have therefore conducted a subgroup analysis and provided additional scenarios for the Committee's consideration of ESK-NS in a later line position in the treatment pathway, i.e., used at 4th line after patients failed 3 or more oral antidepressant (OAD) treatments. This cohort of patients are more resistant and difficult to treat, having failed 3 lines of OAD (5) and the clinical evidence suggests even greater relative efficacy and cost effectiveness of ESK-NS. This is where the higher unmet need in TRD is and therefore ESK-NS has the potential to provide an even more valuable treatment option versus currently available options.

In terms of addressing the Committee's clinical concerns, the EMA, the FDA and other regulators around the world have all concluded that ESK-NS is a clinically effective option and has an appropriate risk-benefit profile for the treatment of TRD. We note that many of these concerns have been examined by the regulators previously. It is important to provide the context in which regulators have considered the ESK-NS trials, which includes some unique challenges in conducting clinical research in mental health, which the Committee may also wish to consider. It must be noted that developing new medicines for the treatment of Central Nervous System diseases is challenging, with a 15% overall probability of

success in clinical trials (6). Approximately 50% of short-term, randomised, controlled trials for approved antidepressants may still fail to show a statistically significant effect (7), primarily due to high placebo responses, which are particularly prominent in mental health trials as opposed to somatic medicine.

Despite these challenges, ESK-NS demonstrated superiority to an active comparator in a population with severe disease that did not respond to at least two previous treatments. The issues relating to the clinical data were also debated by the CHMP, who ultimately judged that the short- and long-term efficacy of ESK-NS in patients with TRD had been established. In an effort to provide a submission package that addresses the specific requirements of the evaluation of a new technology in mental health, Janssen sought NICE Early Scientific Advice (in 2013) and modelling advice from NICE PRIMA (in 2018). It is important for the Committee to note in their consideration of the evidence that we have tried to implement as many of the recommendations from the NICE advice as we have been given. Our response intends to address the additional clinical concerns from the Committee and provide reassurance of the clinical benefit that ESK-NS brings, which was also clearly articulated by the patient representative during the second NICE Appraisal Committee meeting.

A summary of these key issues is provided below, together with the main response points per issue:

Key point 1: Regulatory authorities assessed the ESK-NS clinical data (TRANSFORM-2 and SUSTAIN-1) and concluded they are robust and demonstrate the clinical value of ESK-NS

- The clinical data for ESK-NS should be considered in the wider context of the unique challenges of conducting clinical trials in this therapeutic area. Regulatory agencies approved ESK-NS having discussed similar clinical points as raised to NICE by a small number of clinical stakeholders.
- The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcomes for patients with TRD.
- TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding.
- Response and remission are established and appropriate outcomes for MDD and TRD.
- The four-week duration of TRANSFORM-2 is appropriate and is aligned with clinical trial design guidance from the CHMP.
- No further conclusions can be drawn about the proportion in response in both study arms if the duration of the trial would have been longer.
- The randomised withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results.
- Patients with suicidal ideation were not excluded from ESK-NS trials and patients with high risk of suicide were studied in a separate clinical development program for ESK-NS.
- A robust risk management plan has been agreed with the MHRA.
- Results from an additional long term safety study show no unexpected safety signals.

Key point 2: The outcomes predicted by the economic model are reflective of the outcomes that patients with TRD experience in the long term and the proportion of patients in MDE health state is appropriate, especially when a revised method for subsequent treatments is incorporated.

- The health states used in the model are appropriate, established and based on previous depression models accepted by NICE.
- Results from a targeted literature review of patients with TRD show that long term outcomes of patients with TRD are poor.
- There are limited data to inform the long-term outcomes of patients with TRD. The Fekadu study cohort is not appropriate to compare to the outcome of the economic model, primarily

because it only included a specific subgroup of patients who were discharged following intense multi-modal inpatient hospital treatment. Instead, wider literature from the targeted literature review provides more appropriate and generalisable data on the long-term outcomes of patients with TRD.

- Given the Committee's concerns with the proportion of people in the MDE state, a revised method for subsequent treatments is proposed which reduces the proportion of patients in the MDE state over time. With the revised method for subsequent treatments, the model better represents the long-term outcomes as per the available literature.

Key point 3: The use of the base case MDE utility is appropriate, and an alternative approach which addresses the Committee's concerns using an amended criteria for MDE and a different utility value is provided for consideration

- The use of the MDE utility is appropriate as it represents the quality of life of patients with TRD with moderate-severe disease.
- Given the Committee's concerns with the different thresholds used in the model and clinical trial programme, we provide a scenario which amends the criteria of the MDE health state. This means an alternative source from a subgroup of moderate to severe patients is used to inform the utility of the MDE health state, rather than using the TRANSFORM-2 baseline MDE utility.
- This change in criteria of the MDE health state to represent the moderate to severe MDE would cover the relapse threshold (≥22) used in SUSTAIN-1, and as such addresses the concerns with the different thresholds used in the clinical trial programme and the model.

Key Point 4: It is appropriate to assume different healthcare resource use costs per health state, which could lead to different medical costs between treatment arms. We propose a sensitivity analysis with reduced cost differences among health states to address the Committee's concerns and include additional costs that may be associated with commissioning ESK-NS.

- It is not appropriate to use SUSTAIN-1 to inform HCRU per treatment arm as it was not designed to collect resource use data, cannot provide any conclusions due to a small number of events, does not consider the full cohort of patients with TRD and is not generalisable to resource use in UK clinical practice. Evidence shows that differential costs per health state are appropriate. This has also been the approach used in previous NICE decision making.
- Given the Committee's concerns with the MDE health state, we provide a sensitivity analysis where the costs of the health states are adjusted using the lower bound of the 95% CI costs, resulting in reduced cost differences among health states.
- A sensitivity analysis, with revised health state costs, is provided to address the Committee's concerns, such that it is acceptable to have differential costs per health state.
- Estimates of the costs of commissioning are incorporated into the model and have very minimal impact on the cost effectiveness.

Key Point 5: Revised post ACD-2 scenarios are provided for consideration (full label population), and in response to clinical consultation feedback to NICE on the proposed later line positioning of ESK-NS, we have conducted analyses for this subgroup of patients.

- Given the comments in the ACD, we propose scenarios with 1 later line positioning (nonresponse to at least 3 prior treatments in the current episode) based on the available data of ESK-NS for the Committee's consideration.

[redacted]

Overall summary

Patients who develop TRD have very limited treatment options, and there is a significant unmet clinical need for new and effective medicines that can improve patients' outcomes and experience. The availability of new treatment options in this context has been hindered by the common challenge of developing new medicines and demonstrating superiority against the current standard of care in mental illness, including depression.

ESK-NS, the first medicine with a new mechanism of action approved for patients with TRD, has proven to be effective and generally well tolerated in this difficult-to-treat patient population through a robust development programme. This clinical development programme, which was co-designed with regulatory and HTA authorities, demonstrated robust trial results in this challenging disease area, despite noticeably high efficacy observed in the active comparator arm.

In addition to the issues on the clinical data, we note that the Committee has revised the assumptions in the economic model resulting in a base case ACD 2 ICER of £64,554-£72,158 (at list price). The revised model assumptions in the ACD relate to using the ERG's original approach for subsequent treatments and the assumption of equal non-drug medical costs for patients on ESK-NS and patients on standard of care given the uncertainty on the long-term outcomes predicted by the model. Based on the clinical literature, we have validated the model outcomes and proposed a revised method for subsequent treatments to address the Committee's concerns on the long-term model outcomes that should allow to the Committee to consider different healthcare resource use per health state. The Committees preferred assumption to assume equal non-drug medical costs for both ESK-NS and standard of care is not aligned to the evidence presented and a scenario with reduced cost differences among health states using the 95% CI lower bound health state costs is presented instead. An alternative source for the MDE utility is considered by amending the criteria for the MDE health state to incorporate all of the different thresholds used in the clinical trial programme. Implementing these changes, as well as adding a one-off cost of commissioning to the Committee preferred ACD base case, results in an ICER of [redacted] (including carer disutility) to [redacted] (excluding carer disutility), which shows ESK-NS is a cost-effective use of NHS resources for the full licensed population.

[redacted]

Feedback from clinical experts and consultees have suggested that ESK-NS will mainly be used after non-response to three prior antidepressant treatments. This is largely due to system capacity and convenience of administration. Given this feedback, a scenario considering the clinical evidence for the later line (at least 3 prior failures) positioning of ESK-NS results in an ICER (including updated PAS) of [redacted]x (Including carer disutility) to [redacted] (excluding carer disutility), [redacted]. Given the significant societal burden of TRD, it is important to note that all ICERs presented are an underestimate of the true cost-effectiveness of ESK-NS. We hope that the information that we are providing here will help the Committee to conclude that ESK-NS is a cost-effective use of NHS resources, and that healthcare professionals and patients with clinical warranted need can access this innovative treatment.

1.0 Regulatory authorities assessed the ESK-NS clinical data (TRANSFORM-2 and SUSTAIN-1) and concluded they are robust and demonstrate the clinical value of ESK-NS

Response to ACD sections 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14, 3.16

Key point 1 summary

- The clinical data for ESK-NS should be considered in the wider context of the unique challenges of conducting clinical trials in this therapeutic area. Regulatory agencies approved ESK-NS having discussed similar clinical points as raised to NICE by a small number of clinical stakeholders.
- The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcomes for patients with TRD.
- TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding.
- Response and remission are established and appropriate outcomes for MDD and TRD.
- The four-week duration of TRANSFORM-2 is appropriate and is aligned with clinical trial design guidance from the CHMP.
- No further conclusions can be drawn about the proportion in response in both study arms if the duration of the trial would have been longer.
- The randomised withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results.
- Patients with suicidal ideation were not excluded from ESK-NS trials and patients with high risk of suicide were studied in a separate clinical development program for ESK-NS.
- A robust risk management plan has been agreed with the MHRA.
- Results from an additional long term safety study show no unexpected safety signals.

The results of the ESK-NS clinical trial programme should be considered within the context of clinical trials in depression. There are specific difficulties in conducting trials in this therapeutic area, as recognised by the EMA (8). Analysis of regulatory trials show that CNS drugs are significantly more likely to fail in phase 3 than non-CNS drugs. Approximately 50% of short-term, randomised, controlled trials of approved antidepressants compared to placebo have failed to show a statistically significant effect, primarily due to high observed placebo effects (7). Despite these disease-specific challenges, the European Committee for Medicinal Products for Human Use stated in its public assessment report that '*Overall, the clinical program can be considered comprehensive and supports the use of esketamine as an adjunctive treatment administered concomitantly with a newly initiated oral SSRI or SNRI*'. Of note the FDA's Benefit-Risk Integrated Assessment also highlighted that '*as esketamine is the first drug in a new class of antidepressants, it is important to put its treatment effect into perspective… and the ability to detect even a nominally significant treatment difference by Day 2 sets this drug apart from other antidepressants*'.

Janssen note that the specific concerns raised by some consultees to NICE on the robustness of the clinical data have been raised with the EMA and the MHRA previously. No new data or information was presented to NICE in the 1st ACD consultation that was not already presented to the EMA and MHRA. After considering these issues, both regulatory bodies approved the use of ESK-NS. Given explicit statements in the ACD, however, references to a number of statements which require responses are

addressed below. It is noted that similar points have been addressed in the literature previously (9). Additional responses to other clinical statements made in the ACD are made in Appendix A.

1.1 The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcome for patients with TRD

NICE ACD Section 3.9:

'TRANSFORM-2 measured a statistically significant difference between esketamine nasal spray with newly started oral antidepressant compared with oral antidepressant with placebo after 28 days. The reduction in MADRS score from baseline was 21 for esketamine and 17 for placebo. The committee noted a separation of treatment effect after 2 days (or 1 treatment), which remained for the duration of the 4 weeks. The committee considered that this may not be a true effect on depressive symptoms. A consultee commented that the 4-week duration of the trial has 'little bearing on the treatment for depression'. The committee noted that the NICE guideline on depression recommended an initial assessment at 2 to 4 weeks to assess symptom response to oral antidepressant, but further regular assessments and dose optimisation would be considered after this point. The committee considered that the data still showed a downward trend in MADRS score, with no evidence of flattening, so 4 weeks was not an appropriate endpoint for measuring response and remission for both treatments. Also, a consultee commented that splitting data into 2 groups, response or remission and no response or remission, can lead to an overestimation of differences between arms. The committee acknowledged that splitting the data into 2 groups could have inflated the differences between arms, particularly because the mean reduction in MADRS was near to the threshold for response in both arms at day 28. So, people could meet the criterion for symptom response in 1 arm but only have minimal differences in MADRS score in the other arm. The committee concluded the response and remission evidence from TRANSFORM-2 should be considered with caution because of the duration of the trial.'

Janssen response:

The safety and efficacy of ESK-NS are supported by 19 Phase 1, 4 Phase 2 and 5 Phase 3 clinical studies, and was designed in accordance with the ICH, GCP, and CHMP guidelines. Throughout the development programme, the company sought scientific and regulatory input from the major Health Authorities. The programme incorporated detailed EMA scientific advice, input from the US FDA, and the design of the trials were informed by NICE Early Scientific Advice. The robustness of the design is reflected in the EPAR, which states that '*the clinical development programme can be considered as comprehensive'* and *'innovative'*.

TRANSFORM-2 was the first ever trial in which a novel antidepressant proved clinically and statistically significantly more effective than a newly initiated conventional antidepressant. In TRANSFORM-2, the mean change in MADRS total score from baseline to the end of induction was –21.4 for ESK-NS plus a newly initiated oral antidepressant (ESK-NS + OAD) versus –17.0 for a newly initiated oral antidepressant plus placebo nasal spray (OAD + PBO-NS) (p=0.020). The randomised clinical trials in the clinical programme used a 2-sided 0.05 significance level to establish efficacy.

The minimum clinically important difference (MCID) when groups are compared to each other at the conclusion of a trial is consistently defined as a difference of 2.0 or sometimes even lower (1.6–1.9) in MADRS scores in the literature (10, 11, 12). The threshold of a 2.0 difference in MADRS between study arms has been used as a criterion for establishing clinically meaningful benefit by European Health Authorities (13). Using currently accepted definitions, therefore, a MADRS reduction of 4.0 points between treatment arms in the TRANSFORM-2 study is clinically meaningful. This was achieved despite the higher-than-expected effect observed in the active comparator arm, as acknowledged in the ACD (Section 3.15), which is unlikely to be seen in NHS clinical practice (see Section B.2.3.7 of Form B). The FDA notes that the magnitude of improvement on the MADRS (primary endpoint) is similar to that achieved by other approved antidepressants that were compared only to placebo (14) in a MDD

population, despite evaluation in a TRD study population with greater illness severity and treatment resistance. The FDA's review additionally refers to secondary patient-reported depression [the 9-item Patient Health Questionnaire (PHQ-9)] and functional outcome [the Sheehan Disability Scale (SDS)] measures, which also indicated consistent benefit with ESK-NS + OAD over OAD + PBO-NS providing further support for the clinical meaningfulness of the efficacy findings.

In addition to TRANSFORM-2, consideration of the evidence from other trials also supports the short-term efficacy of ESK-NS. All 3 short-term efficacy studies demonstrated between-group treatment differences of ≥2.0 points change in MADRS total score. The improvement (change from baseline in MADRS total score) seen with ESK-NS was consistent across the Phase 2 and 3 studies, with a similar magnitude of improvement seen in the ESK-NS in the adjunctive Phase 2 studies and the Phase 3 studies with a new OAD. The magnitude of improvement in the comparator arm, however, was higher in the Phase 3 studies leading to the smaller treatment differences (Table 1 below). The results from TRANSFORM-2 are remarkably consistent with the TRANSFORM-1 study.

Table 1: Primary Efficacy Results for Change in MADRS Total Score in Phase 2 and 3 TRD Studies

Study	Treatment Group	Number of Patients	Mean Baseline Score (SD)	LS Mean Change from Baseline to Endpoint (SE)	LS Mean Difference (95% CI)[†]	P-value
TRD2001 (MMRM Day 2)	ESK 0.20mg/kg	9	33.1 (3.55)	-16.8 (3.00)	80% CI -12.9 (-17.93, -7.94)	1-sided p-value 0.001
	ESK 0.40mg/kg	11	33.7 (5.82)	-16.9 (2.61)	-13.1 (-17.93, -8.25)	0.001
	Placebo	10	33.9 (4.15)	-3.8 (2.97)		
TRD2003 Panel A Period 1 (MMRM Day 8)					LS Mean (SE)	1-sided p-value
	ESK 28 mg	11	31.3 (3.80)	-9.8 (2.72)	-5.0 (2.99)	0.051
	ESK 56 mg	11	33.2(6.26)	-12.4 (2.66)	-7.6 (2.91)	0.006
	ESK 84 mg	12	35 (4.22)	-15.3 (2.56)	-10.5 (2.79)	<0.001
	Placebo	33	35 (5.18)	-4.9 (1.74)		
TRD3001 (MMRM Day 28)						2-sided p-value
	ESK 56 mg + oral AD	115	37.4 (4.8)	-18.8 (1.3)	-4.1 (-7.7, -0.5)[#]	0.027
	ESK 84 mg + oral AD	114	37.8 (5.6)	-18.5 (1.3)	-3.2 (-6.9,0.5)[#]	0.088
	Oral AD + Placebo nasal spray	113	37.5 (6.2)	-14.8 (1.3)		
TRD3002 (MMRM Day 28)	ESK (56 mg or 84 mg) + oral AD	114	37.0 (5.7)	-19.8 (1.3)	-4.0 (-7.3, -0.6)	0.020
	Oral AD + Placebo nasal spray	109	37.3 (5.7)	-15.8 (1.3)		
TRD3005 (MMRM Day 28) (≥65 years)	ESK (28 mg, 56 mg or 84 mg) + Oral AD	72	35.5 (5.9)	-10.2 (1.5)	-3.6 (-7.2, 0.07)#	0.059
	Oral AD + Placebo nasal spray	65	34.8 (6.4)	-6.2 (1.5)		

SD=standard deviation; SE=standard error; LS Mean=least-squares mean; CI=confidence interval; AD=antidepressant

† Difference between treatment groups in least-squares mean change from baseline

Median unbiased estimate (ie, weighted combination of the LS mean difference from oral AD + placebo nasal spray), and 95% flexible confidence interval.

In recognition of this consistent effect, the EPAR stated that the placebo-adjusted effect size in the change from baseline to 4 weeks observed in the short-term phase 3 studies can be considered as '*clearly clinically relevant*' (15). In addition, examination of the heterogeneity statistics in the meta-analysis of the three ESK-NS Phase 3 studies reveals an absence of heterogeneity (I^2= 0%). This shows the consistency and robustness of findings across the three trials.

Subsequent to the positive regulatory acceptance of ESK-NS by EMA and the FDA, a number of supportive publications show the clinical benefit and demand for ESK-NS. It is recognised that the ESK-NS acute trials in adults, when pooled, showed a clinically significant difference in MADRS compared to the active comparator arm (-3.84, 95% CIs -6.29, -1.39) (16). In addition, independent meta-analyses have shown the beneficial effect size compared to augmented antidepressant therapies. A recently published meta-analysis, comprising a total of 25 RCTs (26 relevant study arms) with altogether 9004 MDD patients, showed a higher mean difference versus the active comparator arm for the pooled ESK-NS trials (mean difference = 4.09, 95% confidence interval: 2.01 to 6.17) than for the pooled second-generation antipsychotic augmentation trials (mean difference = 2.05, 95% confidence interval: 1.51 to

2.59). The effect size was nearly twice as high versus antidepressant augmentation with second-generation antipsychotics (17).

The TRANSFORM studies have shown the NNT to be 7 and 9 for response and remission respectively, as well as the additional clinical benefits that ESK-NS can bring to patients with TRD (18, 19, 20, 21). Papakostas et al, 2020 (22) also demonstrated the high efficacy of ESK-NS compared to augmentation in a meta-analysis. Comparing the effect sizes and number needed to treat (NNT), the treatment effect for ESK-NS is larger than those of standard OADs or antidepressant augmentation with atypical antipsychotics in TRD (22).

The TRANSFORM-2 results therefore clearly indicate a statistically significant and clinically relevant treatment effect and outcomes for patients with TRD and should be appropriate for Committee decision making.

1.2 TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding.

NICE ACD Section 3.10:

'TRANSFORM-2 showed a 4-point difference between treatment arms on the MADRS scale (see section 3.9). A consultee commented that this was not a clinically significant difference because a minimally improved score of 7 to 9 would be expected to establish clinical benefit for an individual person. The clinical expert commented that for a population in a trial, a mean difference of 4 was clinically significant. The treatment effect of the control arm was greater than would be expected in other trials in depression (see section 3.15). Also, the mean 4-point difference in MADRS score was much smaller than the total effect of the placebo and antidepressant arm, which saw a reduction in MADRS score of 17. The committee noted that there is debate about what is considered a minimal clinically significant difference in the literature. The committee considered that it is difficult to distinguish the following issues from the true difference in treatment effect:

- *the effect of starting a new oral antidepressant at the same time as esketamine*
- *the trial designs and inclusion criteria leading to a much higher placebo response than would be expected (see section 3.15) which could affect relative treatment effect*
- *the non-linearity of MADRS (see section 3.8)*
- *a likely regression to the mean because patients were recruited during the peak of a depressive episode*
- *early 4-week assessment of outcomes (see section 3.9)*

The committee considered there to be differing opinions on the importance of the observed difference but noted the European Medicines Agency (EMA) considered the effect size to be clinically significant. However, the committee also noted that all the TRANSFORM studies used a difference in MADRS score of 6.5 in the power calculations used to estimate sample sizes. The committee concluded that TRANSFORM-2 was not powered to detect a difference of 4 points on MADRS and could potentially have shown a false positive result.'

Janssen response:

The 6.5-point value noted in the ACD refers to the estimated difference used in the power calculations to estimate the sample size for the short-term Phase 3 trials. This difference was based on results from a Phase 2 study where a failing OAD was continued and was ambitious given that a new OAD was not initiated in Phase 2 but was initiated in the Phase 3 studies. In the Phase 2 study, the estimated mean difference change in MADRS total score between the ESK-NS + OAD groups and the placebo + OAD groups were 7.3 for 56 mg and 10.5 for 84 mg. This is in contrast to a treatment difference of 4.0 points on the change in MADRS total scores seen in TRANSFORM-2 study. It is important to note that this difference was statistically significant despite a sample size based on a difference on 6.5 on the MADRS total score. The impact of the difference between the sample size power calculations and the observed results was that the study was underpowered (i.e.., lower chance of demonstrating a statistically

significant difference between treatments based on the assumption of a 6.5-point treatment difference). This could result in an increased risk for the study to show a false negative result (inflation of the type II error because of reduced power). The risk of a false positive result in the TRANSFORM-2 study was mitigated in the study design by the use of the original type I error's alpha. The same type 1 error alpha was used in the sample size calculation and the analysis of the observed data. The type I error was controlled at the 2-sided 0.05 level. In summary, the risk of a false positive was adequately controlled for by the significance levels used in the study.

There is also consistency in the treatment effect of ESK-NS in terms of LS mean change (see Table 1 above), which supports the conclusion there is no false positive result. ESK-NS performs consistently in all Phase 3 studies and the effect is also consistent with Phase 2 results. This is highly unlikely to be due to chance. The difference in the Phase 3 study to the Phase 2 study is driven by the efficacy in the active comparator arm (see Table 1 above), therefore consideration of a false positive effect is inappropriate. The initiation of a new OAD in Phase 3 can be assumed to positively impact the outcomes in the comparator arm, thereby making it more difficult to show a differential treatment effect. Additionally, as noted above in Table 1, the results from TRANSFORM-2 are remarkably consistent with the TRANSFORM-1 study making it less likely that TRANSFORM-2 shows a false positive result.

As such TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding. Janssen believes the TRANSFORM-2 study provides robust evidence of the clinical benefit of ESK-NS, as recognised by regulatory agencies around the world.

1.3 Response and remission are established and appropriate outcomes in depression, and the four-week duration of TRANSFORM-2 is appropriate and is aligned with clinical trial design guidance from the CHMP

NICE ACD Section 3.9:

'TRANSFORM-2 measured a statistically significant difference between esketamine nasal spray with newly started oral antidepressant compared with oral antidepressant with placebo after 28 days. The reduction in MADRS score from baseline was 21 for esketamine and 17 for placebo. The committee noted a separation of treatment effect after 2 days (or 1 treatment), which remained for the duration of the 4 weeks. The committee considered that this may not be a true effect on depressive symptoms. A consultee commented that the 4-week duration of the trial has 'little bearing on the treatment for depression'. The committee noted that the NICE guideline on depression recommended an initial assessment at 2 to 4 weeks to assess symptom response to oral antidepressant, but further regular assessments and dose optimisation would be considered after this point. The committee considered that the data still showed a downward trend in MADRS score, with no evidence of flattening, so 4 weeks was not an appropriate endpoint for measuring response and remission for both treatments. Also, a consultee commented that splitting data into 2 groups, response or remission and no response or remission, can lead to an overestimation of differences between arms. The committee acknowledged that splitting the data into 2 groups could have inflated the differences between arms, particularly because the mean reduction in MADRS was near to the threshold for response in both arms at day 28. So, people could meet the criterion for symptom response in 1 arm but only have minimal differences in MADRS score in the other arm. The committee concluded the response and remission evidence from TRANSFORM-2 should be considered with caution because of the duration of the trial.'

Janssen response:

This section outlines the reasons for why response and remission are appropriate outcomes in depression, and how the four-week duration of TRANSFORM-2 is appropriate to capture the relative clinical benefit of ESK-NS.

The use of and definition of the response and remission outcomes used in TRANSFORM-2 are established, reliable, validated, and appropriate in the field of depression. As noted in the EPAR for ESK-NS, the endpoints (primary and secondary) used for the evaluation of the efficacy of ESK-NS in TRD are considered reliable, validated and are referenced in treatment and development guidelines (23). They have been used throughout many years in clinical practice and hence are appropriate to estimate the differences in treatments. The EPAR states:

'The selection of the endpoints, the measurement of MADRS change from baseline to end point (after 4 weeks in double-blind induction phase) and the difference of this change between treatments is considered appropriate and in accordance with the current guidelines, available literature and clinical practice'.

The CHMP Guideline for Depression products notes that in MDD, a 50% improvement of a patient on a usual rating scale is accepted as a clinically relevant response and that other definitions of responder may be used, e.g., proportion of patients with full remission. This is aligned to the outcomes used in the clinical trial programme and economic model.

This is acknowledged by the EMA, and the EPAR states: *'Besides the differences observed in the change of MADRS from baseline to day28/endpoint between esketamine + oral AD and oral AD + intranasal placebo, responder and remitter rates are also important to demonstrate efficacy in depression'.*

Previously, in NICE TA367 (Vortioxetine for treating major depressive episodes), the Committee agreed that response and remission outcomes were more useful than change from baseline in depressive severity scores:

TA367 FAD, p40: *'The Committee agreed that achieving remission and avoiding relapse were much more useful outcomes than the mean change in a person's depressive symptom severity score for measuring success of treatment in clinical practice.'*

As such, response and remission are considered a relevant outcome as they are established in depression by regulatory bodies and have also been considered by the Committee in NICE TA367 to be relevant.

The 4-week duration of TRANSFORM-2 is clinically appropriate and is of sufficient duration to capture the response expected from the treatment arms, as previously mentioned in the response to Technical Engagement (p.14), This duration was chosen to provide sufficient time for the judgement of onset of efficacy in the OAD + PBO-NS group. Furthermore, a fixed titration to the maximum tolerated dose was implemented to ensure optimal response for the OAD.

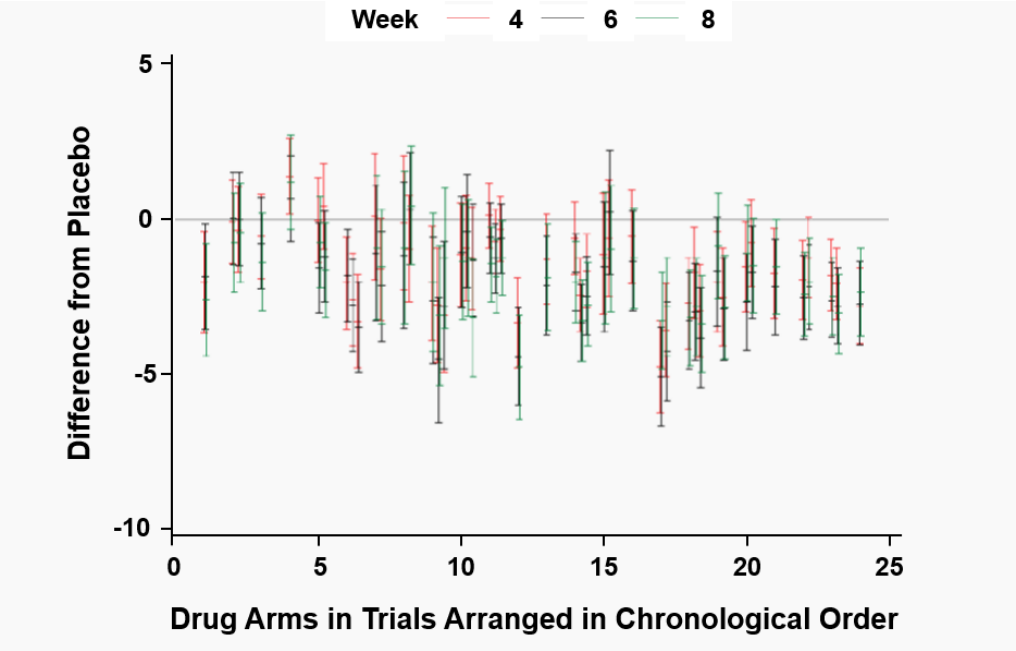
A 4-week study duration is within the recommendations of the CHMP guideline for clinical trials in depression. As well as the recommendations on the design of clinical trials, changing treatment at 4 weeks is aligned to recommendations in clinical guidelines. Good clinical practice would suggest that after 4 weeks of prospective observation with an ineffective antidepressant, these patients would not continue receiving the same OAD that did not show response and improvement in their condition. This view is also taken by the EMA, who stated in the EPAR that:

'contemporary clinical practice which dictates that after 4 weeks prospective observation with one antidepressant these patients cannot continue receiving the same AD that did not show response and improvement in their condition.'

The time point of four weeks to assess treatment response is additionally supported by the recommendations in the British Association for Psychopharmacology (BAP) evidence-based guidelines for treating depressive disorders (24), the Maudsley Prescribing Guidelines in Psychiatry (25), and recommendations in NICE CG90. NICE CG90 recommends that if there is no, or barely any detectable improvement at 2 to 4 weeks, patients should be followed weekly, and consideration given to changing treatment (and not considering adjusting dose) at 3 to 4 weeks. Efficacy data in TRANSFORM-2 was not captured after 4 weeks if response was not achieved.

We note the Committee's concern that there was still a downward trend in MADRS score, with no evidence of flattening, and that 4 weeks may not be an appropriate endpoint. This was a question that the FDA have previously looked at and they conducted a meta-analysis of data from 24 short-term OAD trials to explore further. In the meta-analysis, the OAD-placebo treatment difference was consistent for trials of 4 to 8 weeks' duration, suggesting that it is plausible to shorten OAD trial duration to 4 weeks from 8 weeks (26, 27). The findings from the meta-analysis from the FDA show there are no consistent differences within each trial arm between 4- 8 weeks measurement duration (see Figure 1). Based on these findings, the FDA concluded that the difference between two study arms expected at 8 weeks to be consistent with the difference at 4 weeks.

Figure 1: Antidepressant Trial Duration. Treatment Effect (95% CI) per Drug Arm (adapted from (27))



Despite the TRANSFORM-2 trial showing a downward trend in change from baseline MADRS in both study arms, it is important to note that both arms contained a newly initiated OAD. On the basis of the FDA meta-analysis showing the relative efficacy of OADs to be consistent between 4 and 8 weeks, it would be expected that the relative treatment efficacy would not change if the duration of the trials would be longer.

In addition, there is a greater difference between the median and mean change from baseline MADRS score for ESK-NS than OAD+PBO-NS (see Table 2). This means that there is a greater positive skew towards more patients responding and remitting in the ESK-NS + OAD arm. If longer follow up were to make any difference, the data suggest that it would show an increase in relative efficacy for ESK-NS + OAD. Therefore, any potential bias from a 4-week versus a 6- or 8-week trial duration would be against ESK-NS + OAD.

Table 2: TRANSFORM 2: MADRS Total score change from baseline to day 28

Change from baseline to day 28	ESK-NS + OAD	OAD + PBO-NS
N	101	100
Mean (SD)	-21.4 (12.32)	-17.0 (13.88)
Median	-24.0	-18.5

Whilst we acknowledge that the overall mean change from baseline MADRS score reduction is near the threshold, this threshold is the mean for a cohort, who have a variety of different changes in MADRS scores. In addition, the mean baseline MADRS score is 37, but the threshold for response is not the same for the whole cohort, given the variety of baseline MADRS scores of patients starting in the study. The mean change from baseline cannot be compared against a threshold for response as the response threshold varies by patient.

For TRANSFORM-2, given the individual patient variability in baseline MADRS score, the corresponding variety of response threshold levels in both arms, as well as the similar downward trend in both study arms, no conclusions can be drawn about the proportion in response levels in both study arms if the duration of the trial would have been longer.

In summary, the definitions of response and remission are appropriate for clinical trials in this disease area. The four-week duration of TRANSFORM-2 is sufficient and appropriate, given guidelines for conducting clinical trials and recommendations from clinical guidelines, and evidence from trials of longer duration. It is not appropriate to draw any conclusions about the response levels in both study arms if the duration of the trial would have been longer. There is a suggestion that a longer duration may have increased the relative effectiveness for ESK-NS + OAD, given the positive skew for ESK-NS + OAD. As recognised by the regulators, the response and remission outcomes in TRANSFORM-2 appropriately capture the relative clinical benefit of ESK-NS.

1.4: The random withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results

ACD: Section 3.11 and 3.12:

'SUSTAIN-1 measured withdrawal of esketamine for a randomised population of people whose depression was in stable response or stable remission. The ERG commented that there was potential for selection bias using these criteria. This is because if esketamine is tolerated participants who have the drug for 16 weeks and do not stop (induction and optimisation phases) stay in the trial by design, which selects people who are less likely to be affected by the treatment burden and do not have adverse events that make them stop treatment. After the optimisation phase, randomised participants stopped having esketamine nasal spray and instead had placebo. All participants continued to have oral antidepressant. A consultee commented that there is potential for functional unblinding with this design because participants randomised to placebo may notice the absence of

psychoactive effects. The consequent negative expectations could increase the chance of relapse. The committee understood from consultation that relapses are highest in the first 4 weeks after stopping an active treatment such as esketamine, and this could be consistent with potential unblinding. The clinical expert commented that the number of relapses could have been overestimated. The committee also noted that people with depression in stable response or remission from the TRANSFORM trials who only had placebo had a lower relapse rate than those who stopped esketamine, although this was not explored fully by the company. The committee concluded that the withdrawal design of SUSTAIN-1 may have biased results in favour of esketamine, if patients were unblinded to what treatment they were having.'

'After the first committee meeting, the committee noted that the company had not provided evidence on the effects of withdrawal from esketamine. At consultation a consultee considered that the potential adverse withdrawal effects of esketamine could have confounded the relapse rates of SUSTAIN-1. This was because MADRS is very similar to scales used to measure withdrawal, such as the Physician Withdrawal Checklist (PWC-20). The company considered that there would be no long-term withdrawal effects of esketamine because at this dose it leaves the body quickly. However, the company also did not use data from SUSTAIN-1 for relapse rate in the oral antidepressant with placebo arm in the economic model to avoid any withdrawal effect (see section 3.21). The clinical expert explained that withdrawal effects of ketamine seen in recreational use are from higher doses. Physical responses, such as sweating and shaking, are not expected at this level of dose. The committee noted that anxiety increased in some participants in SUSTAIN-1, 2 weeks after stopping esketamine for both arms, as measured by the PWC-20. The committee concluded that any withdrawal effect would be difficult to distinguish from a change in depressive symptoms because withdrawal is likely to adversely affect people psychologically, including potential increased feelings of hopelessness (see section 3.2).'

Janssen response:

SUSTAIN-1's randomised withdrawal design was recommended and accepted by many health authorities for relapse prevention trials (28, 23) and the early relapses in the trial were not due to unblinding and withdrawal effect but were likely due to the underlying disease and inclusion of vulnerable patient groups. We believe that the SUSTAIN-1 trial is robust and was discussed and agreed upon with both the FDA and CHMP, and the results from the trial were accepted by both regulatory authorities. The population was necessarily enriched because the study's objective was to assess time-to-relapse in patients who remit or respond to ESK-NS. This is representative of clinical practice, as only responders and remitters on ESK-NS would continue treatment past the induction treatment phase, which is reflective of best practice, the SmPC license wording and treatment discontinuation guidance previously presented to the Committee. If patients do not respond, then patients should not continue treatment with ESK-NS and therefore we would ask the Committee to reconsider their argument that the trial selects people who are less likely to be affected by the treatment burden and do not have adverse events that make them stop treatment.

The early relapses in SUSTAIN-1 were not due to unblinding or a withdrawal effect and are most likely due to the underlying disease and inclusion of vulnerable patient groups. These points have been addressed previously and published (14), with the main points to consider are summarised below. The increased early rates of relapse in OAD + PBO-NS group reflect that in clinical practice, patients with treatment-resistant depression (TRD) generally relapse earlier than those with MDD, due to the increased risk of illness and vulnerability in this TRD patient population (5). As noted by the clinical expert in the 2nd ACM, there is a clear distinction between a withdrawal effect and a relapse. In SUSTAIN-1, although there was a high number of relapses in the first month in those switched to placebo nasal spray, it is unlikely that a pharmacologic withdrawal effect contributed given the ESK-NS short half-life. The initial half-life is approximately 30 minutes, with subsequent, sequential half-lives approximately 2 hours and 11 hours. In addition, clinical data from a post hoc analysis (29) in patients enrolled in SUSTAIN-2, of the Physicians Withdrawal Checklist (PWC-20), including the PWC-Withdrawal Symptoms (PWC-WS) subscale, suggests that stopping ESK-NS after short- or long-term use of ESK-NS is highly unlikely to be associated with withdrawal syndrome, as assessed by stability, frequency, onset, and severity of PWC-WS, SAEs reported during follow up, the low rate of positive urine drug screens and the absence of drug-seeking behaviours.

Relapse rates from the literature also provide helpful context. The high rate of early relapse in the SUSTAIN-1 OAD + PBO-NS arm (45.3%) is similar to that observed after cessation of electroconvulsive therapy (30), of which there are no known rebound effects. In the STAR*D study, relapse rates in patients who failed two or three antidepressant treatments and achieved remission following the 12- to 14-week acute treatment phase were high at 42.9% and 50.0% respectively. The mean times to relapse were 3.9 and 2.5 months, respectively, even while continuing the treatment to which they had responded. There are also significant relapse rates in the maintenance study with quetiapine XR (34.4% relapse rate for the placebo group) (11) and ECT studies (50.0% relapse after 1 year) (31).

The relapse rates of the group of patients who responded to only OAD + PBO-NS in TRANSFORM-2 should not be compared to the group of patients who responded to ESK-NS + OAD in TRANSFORM-2, as there is no basis in randomisation for such a comparison. This was previously addressed in the Company response to the ERG Clarification Questions (B10b, p48). This is also presented in the footnote to Figure 11 in Appendix D.2.2 of the Company Submission, this group of patients within the SUSTAIN-1 trial is not informative on the transition probability of relapse and loss of response on OAD, as these patients did not achieve remission or response while on esketamine and are therefore not included in the efficacy analyses.

Specific study design measures were incorporated to ensure the trials were appropriately double blinded to both the assessors and study participants, as noted in Section B.2.3.3 of the Company Submission. The EPAR states that '*a higher-than-expected response in the oral AD + intranasal placebo arm is an opposite effect than is expected in case of unblinding'*, which shows that patients were not able to distinguish between the active treatment and the placebo nasal spray. This shows that unblinding did not affect the results of SUSTAIN-1.

Janssen conducted additional pre-specified and post-hoc analyses to assess whether potential unblinding could have affected the results. Specifically, Janssen conducted a sensitivity analysis, a patient-level adverse event assessment, and a mediation analysis on TRD3003 data to assess the potential impact of dissociation on treatment effect. The results indicated that the absence or presence of dissociative symptoms did not account for or contribute significantly to the effect of ESK-NS suggesting that unblinding did not exert a major effect on the antidepressant effect difference between groups. Further details on these results can be found in Appendix B.

1.5: Patients with suicidal ideation were not excluded from ESK-NS trials and patients with high risk of suicide were studied in a separate clinical development program.

ACD, Section 3.14:

'The expert from the NICE guideline on depression noted that excluding people with an acute suicide risk reduces the generalisability of the trials because people with treatment-resistant depression are likely to have an increased risk of suicide. A clinical expert also noted that excluding suicide risk was a concern because suicidal ideation is often an integral part of the disease. The committee noted that many people referred to a psychiatrist (a requirement of the SPC) in NHS clinical practice would be at higher risk of suicide'

Janssen response:

Patients with suicidal ideation and hence those patients that have an increased risk of suicide, were not excluded from the studies, as previously noted in Janssen's response to NICE's Draft Technical Report. The prevalence of suicidal ideation and the lifetime history of suicidal behaviour reported in the studied population is consistent with published data (32).

Between 15% and 31% of the patients across the Phase 3 studies in TRD had suicidal ideation at baseline, and 25% to 37% of the enrolled patients had a lifetime history of suicidal ideation. Whilst suicidal ideation is a core component of MDD and TRD, there is only a small overlap between patients at high risk of suicide and patients with TRD. In a database analysis of the CRIS database of the South London and Maudsley (SLaM) region, only [redacted] of patients with TRD were recorded as high-risk of suicide. There is a separate indication of ESK-NS for patients with MDD who have active suicide ideation with intent, for which the main results of the entire development program have been published (2,3). The efficacy of ESK-NS has been demonstrated to be similar in this population. In 2018, ESK-NS received a Promising Innovative Medicine (PIM) designation from the MHRA for this indication. In July 2020, the FDA approved ESK-NS for this second indication and this is currently under review by the EMA, with license expected in Q1 2021.

1.6: A robust risk management plan has been agreed with the MHRA

ACD, Section 3.16:

'The EMA identified some risks of esketamine use in the SPC. These included drug abuse, transient dissociative states and perception disorders, disturbances in consciousness, and increased blood pressure. At the first meeting, a registry was suggested to monitor how much esketamine a person has and to prevent people from getting esketamine from more than one source. The clinical expert also suggested including IV ketamine on this registry for the same reasons. They explained that there is likely to be an increased risk of misuse or abuse in people who are dependent on alcohol and drugs. The NHS commissioning expert explained that, because esketamine is a schedule 2 drug, it is subject to the full controlled drug requirements relating to prescriptions and storage (see section 3.30). The committee acknowledged that the monitoring period would likely mitigate the other risks identified in the risk management plan and the committee did not need to consider these further. However, it noted that a registry must be considered when administering and monitoring esketamine to prevent abuse and misuse. After consultation, the committee received further comments about the safety of esketamine. Namely, that the clinical evidence showed there were 3 suicides in people that stopped esketamine in a population who had no recent suicidal ideation or behaviour. There were no suicides in people who had placebo, although people had placebo for less time. In SUSTAIN-1, there were also a higher number of hospitalisations and clinically relevant events. The committee recognised the numbers reported were very small but enough to doubt that there would be more crisis hospitalisation for placebo than esketamine. It concluded that the precautions in the SPC were appropriate regarding risk of suicide and management through increased monitoring, particularly during early treatment and after dose changes.'

Janssen response:

Janssen acknowledge that drug abuse is included and has been identified in the European Union Risk Management Plan (EU-RMP) as an important risk for ESK-NS and further discussion with the MHRA have led to a controlled access program. To mitigate the risk of abuse both routine and additional risk minimisation measures have been put in place, as follows:

- Routine Risk minimisation measures:
    - SmPC section 4.4,
    - Patient Leaflet Section 2,
    - Administration under the direct supervision of a healthcare professional (SmPC Sections 4.2 and 4.4, PL Section 3, and Instructions for Use),

    - Limited pack sizes,
    - Legal status: Special and restricted medical prescription with categorisation at the Member State level.
- Additional Risk Minimisation Measures:
    - Healthcare Professional Guide,
    - Patient Guide,
    - Controlled Access Program.

The objective of the Controlled Access Program is to implement the most appropriate system for the purposes of minimising the risk of drug abuse at national level. The key elements of this program are:

- ESK-NS is intended to be self-administered by the patient under direct HCP supervision, and should be dispensed to the healthcare settings where administration takes place, as agreed at the Member State level, based on local legal requirements and/or local healthcare systems,
- The setting needed for product administration.

Janssen has tailored the details of the Controlled Access Plan with the MHRA prior to launch of the product in the UK. The MHRA advised that, as part of the Controlled Access plan, a registry matching individual patients with the drug was recommended. Janssen is working closely with the MHRA to finalise the protocol for the registry and shares monthly reports of orders of ESK-NS, with the purpose of a continuous monitoring of the potential for abuse.

Lastly, it has to be noted that the ESK-NS device has been designed with several features to deter misuse/abuse:

- The device is single-use and delivers 2 sprays, with no sprays remaining after the second spray is discharged. The device does not require priming before use.
- Device and primary container are extremely difficult to disassemble (force required to pull the device apart is ~13 pounds), which is a deterrent to disassembly.
- The device has a nominal fill volume of 230 uL and a delivery volume of 200 uL. The average measured residual volume left in nasal spray device after actuation is ~30 uL (4 mg base).

In addition, clinical data as noted above, from a post hoc analysis in patients enrolled in SUSTAIN-2 (NCT02497287), of the Physicians Withdrawal Checklist (PWC-20), including the PWC-Withdrawal Symptoms (PWC-WS) subscale, suggests that stopping ESK-NS after short- or long-term use of ESK-NS is highly unlikely to be associated with withdrawal syndrome; patients in the study also did not experience drug seeking behaviour during the study.

In conclusion, the risk of abuse of ESK-NS is currently addressed by a comprehensive set of measures agreed with, and endorsed by, the EMA and the MHRA and include the proposal for a registry, which will be funded by Janssen.

1.7: Results from an additional long term safety study show no unexpected safety signals:

Interim unpublished data from the long-term safety study SUSTAIN-3 show that there were no new safety concerns identified with continued intermittent ESK-NS dosing of up to 30 months (54 [4.7%] patients) as compared with the already determined safety profile in patients exposed to ESK-NS for up

to one year. Specifically, longer-term exposure to ESK-NS showed no additional concerns related to cognition, suicidality, abuse potential, lower urinary tract symptoms, renal adverse reactions, or hepatic adverse reactions. Further results from this long-term safety trial will be available in Q2 2021 and can be shared with NICE.

Section 1 conclusion

Overall, regulatory authorities assessed the whole body of evidence for ESK-NS and concluded they are robust and demonstrate the clinical value of ESK-NS. The clinical programme of ESK-NS should be considered in the wider context and with regards to the difficulties of conducting clinical trials in this therapeutic area. The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcome for patients with TRD. A SUSTAIN-2 post-hoc analysis showed that stopping ESK-NS after short- or long-term use is highly unlikely to be associated with withdrawal syndrome. The randomised withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results.

2.0: The outcomes predicted by the economic model are reflective of the outcomes that patients with TRD experience in the long term and the proportion of patients in MDE health state is appropriate, especially when a revised method for subsequent treatments is incorporated.

Key point 2 summary

- The health states used in the model are appropriate, established and based on previous depression models accepted by NICE.
- Results from a targeted literature review of patients with TRD show that long term outcomes of patients with TRD are poor.
- There are limited data to inform the long-term outcomes of patients with TRD. The Fekadu study cohort is not appropriate to compare to the outcome of the economic model, primarily because it only included a specific subgroup of patients who were discharged following intense multi-modal inpatient hospital treatment. Instead, wider literature from the targeted literature review provides more appropriate and generalisable data on the long-term outcomes of patients with TRD.
- Given the Committee's concerns with the proportion of people in the MDE state, a revised method for subsequent treatments is proposed which reduces the proportion of patients in the MDE state over time. With the revised method for subsequent treatments, the model better represents the long-term outcomes as per the available literature.

ACD Section 3.17:

*'The company economic model consisted of 5 health states: major depressive episode (MDE), response, remission, recovery and death. The transitions between each health state were determined by the relapse, remission and response rates in TRANSFORM-2, SUSTAIN-1 (see section 3.8) and values in the literature, for example the STAR*D trial (a large-scale clinical trial for people with depression). All people start in the MDE state and the initial treatment uses response and remission data from TRANSFORM-2. This is followed by 3 more potential subsequent treatments after non-response or relapse, and then a non-specified mixture of treatments. The model output suggests that within 1 year, 78% of people with treatment-resistant depression in current clinical practice do not have symptom response to any treatments long-term. So, they then occupy the MDE state for the remainder of the time horizon. At consultation, a consultee stated that improvements in depression are generally maintained at the end of acute treatment, and on average symptoms improve further. Another consultee considered that depression can be highly episodic, with a good success rate when people adhere to treatment. The committee heard that there is minimal long-term outcome data for people with treatment-resistant depression. One study in a tertiary care setting (inpatients) suggested that half of people are in remission at a median of 3 years follow-up. This population would have more severe depression than people with treatment-resistant depression in the clinical evidence. The clinical expert estimated that currently 20% to 30% of people with treatment-resistant depression have chronic longer-term disease that has not responded to any treatment. The committee considered that the economic model likely overestimated the number of people in the MDE health state in both treatment arms. The ERG noted that this was likely because subsequent treatment effects had been underestimated (see section 3.18) and modelling a high relapse and recurrence rate (see section 3.21). The committee also considered that the health states used in the model were not the most appropriate for the economic modelling. This was because it was likely there was heterogeneity of costs and utility within these health states. The committee concluded that the economic model did not reflect the course of the disease and does not reflect the episodic nature of the condition.'*

2.1: The health states used in the model are appropriate, established and based on previous NICE depression models

The health states used in the model are based on the established structure of depression models, (NICE CG90) and demonstrated in the systematic literature review of economic evaluations included in the company submission. The models have been considered appropriate for previous NICE decision making (CG90) in the past and addresses the critique from the NICE Committee in TA367, as noted by NICE PRIMA. This was also noted by the ERG who judged the model structure and health states to be plausible in the ERG report (*p111): 'ERG comment: The model structure seems plausible and responds appropriately to the critique in TA367'.* The Committee notes that the majority of patients will follow an episodic course of the disease and some other patients may have a more chronic longer-term disease, as noted by the clinical expert. Specifically, the transition from the MDE health states to response or remission and back to the MDE health state (relapse) appropriately represents the disease course.

In addition, the current economic model and health states adequately captures the disease progression over time and the long-term model outputs (as shown in Section 2.2). We have further addressed the Committee's concerns around the proportions of patients remaining in the MDE health state below and these are aligned to other studies showing the long-term outcomes of patients with TRD (33, 5). We therefore ask the Committee to reconsider their conclusion that the economic modelling does not reflect the course of the disease, or the episodic nature of the disease given the literature and the precedence in previous NICE decision making.

2.2: Results from a targeted literature review of patients with TRD shows that long term outcomes of patients with TRD are poor

The single study referenced in ACD Section 3.17 (as described by Fekadu et al and Wooderson et al) should be considered with caution, as it was conducted in a tertiary care setting and therefore has limited generalisability to NHS clinical practice and is not consistent with the literature. The results of this study should therefore not be used to validate the long-term outcomes in the TRD economic model. Findings from the wider literature identified through a targeted literature review shows that patients with TRD who do not respond adequately to a series of medications are at a very low likelihood of responding to additional medications, or to augmentation and combination pharmacologic strategies (34, 35, 36). Further details on these points are provided below.

The ACD refers to the study conducted by Wooderson et al when concluding that the proportion of patients in MDE over time are overestimated. The relevance of the Wooderson study is limited when estimating the long-term outcomes of a full patient cohort with TRD in the NHS. The long-term remission and recovery outcomes from this study are not appropriate to compare to the output of the economic model for a number of reasons. Firstly, it is acknowledged in the ACD that the population included in the study is different to the general population of patients with TRD. In the study, the population was severe despite patients being included if they failed at least one prior OAD, which is different to the population of TRD as per the license wording for ESK-NS. Secondly, this study cohort (in treatment conducted in an inpatient setting) does not represent the current NHS clinical practice setting for TRD treatment, which is mainly in an outpatient setting. Furthermore, the mean duration of inpatient admission in this group was 7.2 months, which cannot be considered average current NHS standard of care for patients with TRD. Clinical experts involved in the service provision of this treatment indicated the intensive multidisciplinary inpatient treatment that these patients received is no longer

available within the NHS due to its resource intensity. Finally, rather than informing the long-term outcomes of patients with TRD, the findings from the Wooderson et al show that patients who received prolonged intensive multidisciplinary inpatient treatment, of which a large proportion were in response (48%) or in remission (45%) at discharge, were subsequently associated with improved long-term outcome during follow up. This cohort had high response and remission rates at baseline and therefore should not be used to estimate the proportion of patients who have a treatment response and remission over the long term. A large proportion of these patients maintained their initial clinical improvement after 3 years (median) post-discharge. NICE PRIMA advice summarised this study by stating that these 'long-term outcomes were at the high end of what might be expected clinically and were achieved after intense multi-modal treatment.'

To better reflect the totality of the evidence base on this topic, a targeted literature review was conducted, of which further details are provided in Appendix I. The results show that the long-term outcomes of patients with TRD are poor. Only low rates of response (<10-25%) and sustained remission (≤12%) are achieved in current clinical practice for up to 3.5 years (Table 3). For the minority who achieve benefit, these patients are more likely than patients with treatment-responsive MDD to experience relapse and recurrence and have lower remission rates (5). A summary of the findings from the literature are provided below for the Committee's consideration.

The STAR*D study shows that patients with TRD have poor long-term outcomes, and lower remission rates per each treatment line received. Based on the study, about 90% of patients who do not respond to two adequate treatments will likely require long-term management for a depressive illness, which is unlikely to remit with currently available treatments (37). This study is the most comprehensive prospective study conducted in the field of MDD/TRD and enrolled 4,041 outpatients. The study included patients who met DSM-IV criteria for nonpsychotic major depressive disorder. In TA367, the Committee heard that the STAR*D trial provided the best available data for the prognosis for people having subsequent lines of treatment. Patients were followed through up to four lines of OAD treatment for both MDD and TRD, and those not achieving remission with, or unable to tolerate a treatment line were encouraged to move to the next treatment line. After steps 3 and 4 (whereupon patients would have failed two and three OADs and would therefore be considered to have TRD), remission rates were 13.7% and 13.0%, and response rates were 16.8% and 16.3% respectively. The probability of sustained benefit (achieving remission and maintaining remission) dropped to 4.85% and 3.76% at the higher degrees of treatment resistance in 3rd and 4th line MDD, respectively (38). This is an even lower rate compared to the outputs of the economic model with the original Janssen assumptions. NICE PRIMA also referred to the long-term outcomes identified in the STAR*D publication to inform the model.

NICE PRIMA identified the Dunner et al, 2006 paper (33) when searching the literature for the long-term outcomes of patients with TRD. In a long term prospective observational study, Dunner et al show that patients with TRD had a very low likelihood of sustained treatment response after 2 years. This was despite receiving a variety of treatments consisting of various classes of antidepressants with augmentation and combination strategies, psychotherapy, and ECT. More than 90% of patients did not remit after 2 years of active treatment. Furthermore, as defined by the SF-36, patients experienced a poor quality of life throughout the duration of the study despite receiving active treatment. The cohort showed 3-, 12- and 24-month response rates of 5.8%, 11.6% and 18.4% respectively. For remission, the 3-, 12- and 24-month rates were 1.7%, 3.6% and 7.8% respectively. The majority of patients with a response at any one point during the study had only intermittent and transient response patterns. There were no persistent remitters during the first year of the study (no patients had more than 1 remission visit during the first year). Over the 2 years, 65% of patients had no response and 81% of patients did not go into remission. The study concluded that despite the wider range of treatment options currently

available for depression, the response rates, remission rates, and quality of life show that most patients with TRD continue to have significant symptomatology and functional disability. This finding supports the long-term output of the economic model, because it showed that there is little meaningful, sustained improvement for the vast majority of patients with TRD with standard care over 1 to 2 years.

Additional publications identified by the targeted literature review are also supportive of the papers identified by NICE PRIMA. A cohort of UK patients with TRD similarly had poor long-term outcomes (Fonagy, 2015). Fonagy, 2015 (39) reported the effectiveness of long-term psychoanalytic psychotherapy (LTPP) as an adjunct to treatment as usual (TAU) according to UK national guidelines, compared to TAU alone, in patients with long-standing major depression who had failed at least two different treatments and were considered to have treatment resistant depression. The study used a pragmatic long-term randomised controlled study design. Patients were recruited from primary care from February 2002 to May 2009 and assessed at the Adult Service of the Tavistock & Portman National Health Service (NHS) Foundation Trust in London. LTPP consisted of 60 (50 min) sessions of once-weekly individual psychoanalytic psychotherapy over 18 months. TAU consisted of interventions as directed by the referring practitioner. This could include referral for other specialist provisions. At 18 months, full remission was infrequent, with rates of 9.4% for LTPP + TAU vs 6.5% for TAU. Complete remission was infrequent in both groups at 42-month follow-up (14.9% vs. 4.4% respectively).

Additional findings from a US registry of patients with TRD also show the poor outcomes of treatment as usual (40). The study compared a cohort of patients treated with vagal nerve stimulation (VNS) and TAU with patients receiving TAU alone (i.e., any treatment available to psychiatrists, including ECT). The TAU arm (which is most likely to represent clinical practice) had 12% remission and 25% response after 1 year. It is noted that the definition of remission is a MADRS total score ≤ 9 at any postbaseline visit and did not need to be sustained. The authors explain there may have been differences in baselines illness status or frequency of visits in the registry design of the study to explain the higher response and remission compared to Dunner et al.

Table 3: Results from long term TRD cohorts

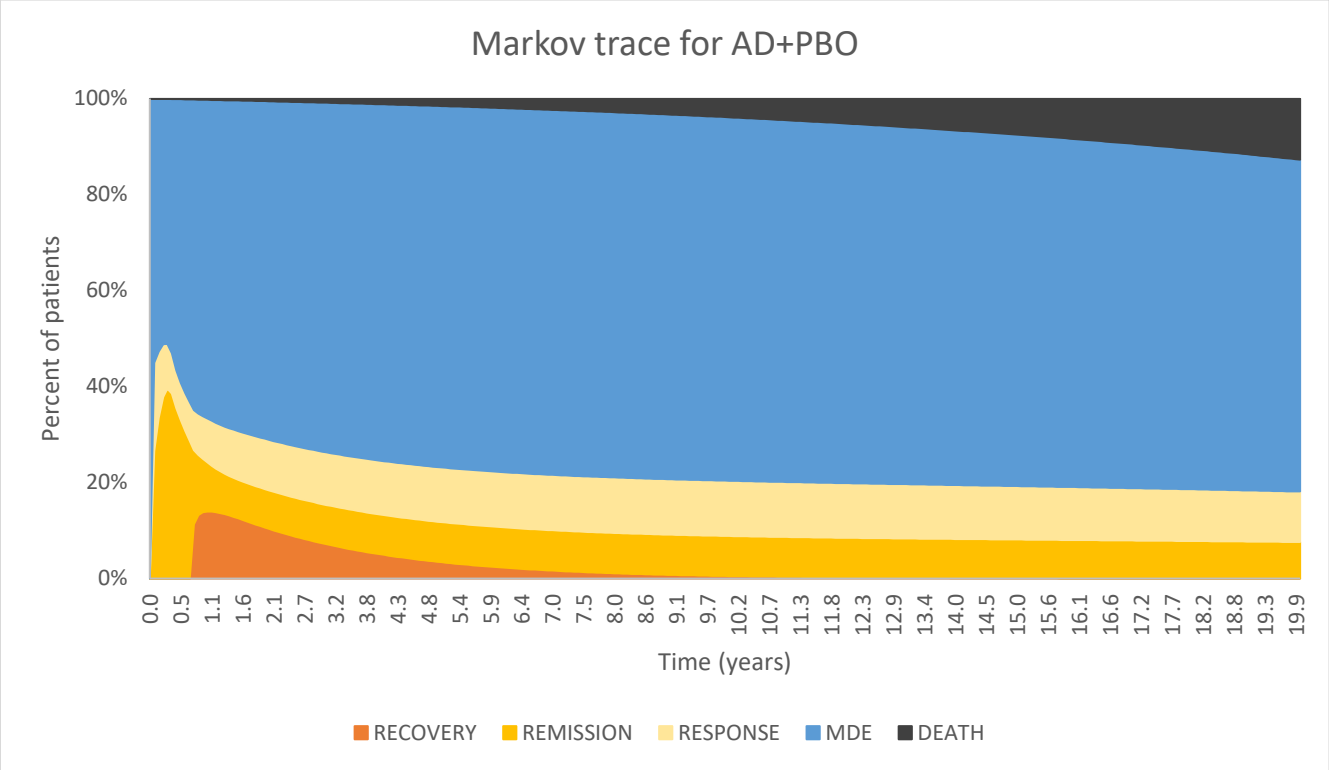
Study	Outcome time point	Definition of remission	Remission on current treatment	Response on current treatment
Rush et al, 2006 - STAR*D (38)	1 year	A QIDS-SR score of ≤5 (equivalent to ≤7 on the HRSD) defined remission	4.85%	Not reported
Dunner et al, 2006 (33)	2 years	IDS-SR-30 score of ≤14	8%	18.4% (including remission; ≥50% decrease in total baseline score, hence including remission)
Fonagy et al, 2015 (39)	1.5 years	HDRS-17 score of 8 or less	6.5%	Not reported
Fonagy et al, 2015 (39)	3.5 years	HDRS-17 score of 8 or less	4.4%	Not reported
Aaronson et al, 2017 (40)	1 year	MADRS total score ≤ 9 at any postbaseline visit	12%	25% (including remission; ≥50% reduction from baseline MADRS score at any postbaseline visit)
Output from Markov model (revised subsequent treatment approach)	1 year	MADRS total score ≤ 12	10.2%*	9.1% (response not remission)
Output from Markov model (revised subsequent treatment approach)	3.5 years	MADRS total score ≤ 12	8.2%*	11.1% (response not remission)

*Recovery is not included in the % remission reported for the Markov model

In addition to the published literature identified in the targeted literature review, data from a recently conducted prospective observational study of patients with TRD are similar. This TRD cohort study was set up to collect data from routine clinical practice in Europe (including the UK), with the aim of understanding the demographics and characteristics of patients with TRD, the current treatment patterns of routine clinical care, and the clinical, social and economic outcomes of routine clinical practice in the treatment of patients with TRD. The cohort of patients, and subgroup of UK patients, demonstrate poor outcomes in current clinical practice. The outcomes of UK patients with TRD are directly relevant for this appraisal, with more than 80% of patients experiencing no response and more than 90% experiencing no remission at 1 year (41). Of the minority of responders at month 6, more than half (52%) lost the response at month 12, meaning that long-term stability of improvement was only present in a minority of patients.

The full evidence shows that the long-term outcomes of patients with TRD are low, with remission estimates ranging after 1 year ranging from 4.4% to 12%. The literature supports the proportion of people in the MDE health state and has shown the proportion not to be overestimated when the wider literature is taken into consideration.

Figure 2: Markov trace using revised base case assumptions (including revised method for subsequent treatments)



The revised method of subsequent treatments (see Section 2.3) results in a Markov trace which can be considered conservative compared to the poor long-term outcomes reported in the literature (Figure 2). At one year, the model estimates that 9.1% of patients are in response, 10.2% are in remission, 14.0% are in recovery, and 66.4% are in MDE (0.3% in death). Therefore, the cost effectiveness estimate for ESK-NS should be considered conservative since the number of patients ending up in the costly MDE state over time might be lower in the model than in clinical practice.

2.3 Given the Committee's concerns with the proportion of people in the MDE state, a revised method for subsequent treatments is proposed which reduces the proportion of patients in the MDE state over time

The Committee considered the economic model overestimated the proportion of patients in the MDE health state, likely due to the subsequent treatment effects. The inputs for the subsequent treatments and best supportive care (BSC) were based on published literature sources (5, 42). We do not consider that the economic model overestimates the proportion of patients in the MDE health state, especially considering findings from the targeted literature review (Section 2.2). Given the discussion at the 2nd ACM, and the clinical expert's view, however, we have provided an alternative scenario for consideration. This uses the method for subsequent treatments as proposed by the ERG in their response to Janssen's first ACD response, which has the impact of reducing the proportion of patients in the MDE health state over time.

The conclusion that the model does not reflect the disease is impacted by the preference for a 20-year time horizon. As mentioned previously, we believe the increase in time horizon increases the uncertainty in the model output. In contrast, a 5-year time horizon, as used in the original company base case, reduces the uncertainty and is aligned to the approach used in other depression models. Given the use of a 20-year time horizon, the long-term outcomes are impacted significantly by the efficacy of subsequent treatments and the efficacy of BSC.

The ERG initially proposed an alternative approach for subsequent treatments in the ERG report. This has been subsequently adopted by the Committee in the ACD 2 base case ICER of £64,554 and £73,158 per QALY. The original ERG approach, as favoured by the Committee, is not appropriate, for the reasons outlined below.

The ERG based the original approach for subsequent treatments on the efficacy observed in the TRANSFORM-2 active comparator arm. As also noted by the Committee, the efficacy in the TRANSFORM-2 active comparator arm was high and is likely to be caused by a '*much higher placebo response than would be expected*'. Thus, the use of the unadjusted OAD arm from TRANSFORM-2 is already conservative for the first line TRD treatment in the model. The high efficacy in the TRANSFORM-2 OAD arm is then amplified in the efficacy of subsequent treatments by using the original ERG approach, see Table 4 below. This is because the original ERG approach results in efficacy values for subsequent treatments which are much higher than the equivalent efficacy observed in patients with TRD (line 3 and line 4) in the STAR*D study, which showed response and remission rates of ~17% and ~13% respectively at 12-14 weeks.

In response to the original ERG approach, Janssen outlined the rationale for why the approach was not appropriate, as explained above. The ERG recognised that the original ERG approach for subsequent treatments had limitations and is likely to be inappropriate (ERG comment on Company ACD Response):

ERG: *'Indeed, at TRD line 2 they are higher than those from STAR*D, particularly for remission. This is largely due to the TRD line 1 values for OAD (not included in Table 6) from TRANSFORM 2 being higher than those from STAR*D, i.e. 26.6% and 18.4% vs. 13.7% and 16.8% respectively for remission and response. Therefore, the ERG have adjusted the values for subsequent treatment by assuming only for the purpose of calculating these values that those at TRD line 1 are those from STAR*D, i.e. 13.7% and 16.8%.'*

Instead of the original ERG method for subsequent treatments, we suggest the Committee should consider the same approach as suggested by the ERG in its response to Janssen's ACD1 response. This revised ERG method also addresses the Committee's concerns with the 4-week adjustment in the base case company approach to modelling subsequent treatments. As shown above, the ERG agrees with this method for estimating the efficacy of subsequent treatments and BSC. Details on the methodology for the revised approach for subsequent treatments can be found in Appendix C. The resulting response and remission values when using this revised ERG method for subsequent treatments, as well as original company base case and original ERG method, are found in below in Table 4.

Table 4: Efficacy of subsequent treatments and BSC phase using different approaches

	Base case company method (after adjustment to 4-week estimates)		Original ERG method		Revised ERG subsequent treatment method (based on ERG comment on Company ACD response)	
	Remission	**Response***	**Remission**	**Response***	**Remission**	**Response***
TRD line 2	3.5%	0.9%	25.2%	17.8%	12.8%%	3.4%
TRD line 3	2.7%	0.7%	23.9%	17.3%	12.1%	3.5%
TRD line 4	2.1%	0.4%	22.7%	16.8%	11.5%	3.7%
BSC (TRD line 5 and all subsequent later lines)	0.4%	0.8%	21.5%	16.3%	10.9%	3.8%

*Response excluding remission

When using the revised ERG subsequent treatment method inputs, the proportion of patients in the MDE health state is reduced. The Markov trace might now be considered more representative of long-term outcomes. Aligned to the literature, the model output in the Markov trace shows that approximately 10% of patients with TRD experience remission at 1 year. Considering the long-term outcomes from the literature, the proportion of people in the MDE health state can be considered appropriate or even conservative.

3.0 The use of the base case MDE utility is appropriate, and an alternative approach which addresses the Committee's concerns using amended criteria for MDE and a different utility value is provided for consideration

Key point 3 summary

- The use of the MDE utility is appropriate as it represents the quality of life of patients with TRD with moderate-severe disease.
- Given the Committee's concerns with the different thresholds used in the model and clinical trial programme, we provide a scenario which amends the criteria of the MDE health state. This means an alternative source from a subgroup of moderate to severe patients is used to inform the utility of the MDE health state, rather than using the TRANSFORM-2 baseline MDE utility.
- This change in criteria of the MDE health state to represent the moderate to severe MDE would cover the relapse threshold (≥22) used in SUSTAIN-1, and as such addresses the concerns with the different thresholds used in the clinical trial programme and the model.

ACD, Section 3.8:

'The Montgomery-Asberg Depression Rating Scale (MADRS) measures severity of depression. It is scored between 0 and 60, 0 meaning no depressive symptoms. Primary outcomes of response and remission in TRANSFORM-2 and relapse rates in SUSTAIN-1 were measured using MADRS. Moderate to severe depression was defined in TRANSFORM-2 as a MADRS score of 28 or more and the mean baseline MADRS score of the participants was around 37. Symptom response was defined as a reduction in score of 50% or more from baseline. The clinical expert explained that this is a standard criterion for response. Remission was defined as a MADRS score of 12 or less with minimal or no symptoms. The clinical expert considered that remission is normally measured by a MADRS score of 10 or less (as in NICE technology appraisal guidance on vortioxetine) but that this would not substantially affect the results. Relapse was defined as a MADRS score of 22 or more for 2 consecutive assessments or other clinically relevant event such as hospitalisation for depression. Recovery was defined as symptoms remaining in remission for about 9 months and recurrence was defined as depression relapsing after recovery. The clinical expert noted that MADRS is non-linear, meaning that a change in score at the lower end of the scale does not mean the same, in terms of clinical importance, as a change in score at higher end of the scale. The committee noted that remission and relapse are fixed to MADRS, but response measurement depends on the score at baseline, which complicates interpretation. The committee also noted that the score used for relapse was not equivalent to the MADRS score for moderate to severe depression, which affected the health state utility values and transitions in the economic model (see section 3.21 and section 3.23). The committee took this into account in its decision-making.'

Janssen Response:

The use of the baseline MDE utility from TRANSFORM-2 throughout the model is appropriate because it reflects the utility of patients in a moderate-severe depressive health state. In the model, the transition from response or remission to MDE represents a relapse. Relapse in MDD is defined as a return of depressive symptoms after effective acute treatment with an antidepressant therapy (43). The relapse criteria in SUSTAIN-1 for MADRS required that a patient met a MADRS score of 22 or higher on 2 assessments. A MADRS score of 22 or higher is consistent with presence of MDE symptoms of moderate severity, with a score of 20-34 being regarded as moderate severity. This was done to ensure that the event was a true relapse and not related to a temporary fluctuation. Although the relapse threshold used in SUSTAIN-1 was based on relapse criteria used in other antidepressant maintenance studies (44), some maintenance studies have used a lower threshold of ≥18 (45). As such, a score of 22 or higher is considered an adequate threshold for defining relapse after established remission.

In SUSTAIN-1, the observed average MADRS score of patients who relapsed also fulfils the definition of the MDE health state used in the model. [redacted]. This means that the majority of patients who relapsed fulfilled the criteria for the MDE health state, and the mean MADRS was [redacted] than the threshold of 22 used in SUSTAIN-1.

The MDE utility used in the base case approach is appropriate to apply for patients who experience relapse, and also for patients who do not respond to any treatment and remain in their initial MDE health state. There are multiple reasons for this:

1. The use of the baseline TRANSFORM-2 utility directly reflects the definition of the MDE health state of patients with TRD. The data are taken directly from patients with TRD, and as such match the definition of the MDE health state.

2. Whilst TRANSFORM-2 include patients with a MADRS total score of ≥28, other studies that also include UK patients, show similar utilities, and included patients with a MADRS total score of ≥20. This covers the thresholds used in the SUSTAIN-1 study for relapse. Results of an observational study in a TRD population (46) that included UK patients shows utility similar to TRANSFORM-2. The mean utility score from 243 patients with TRD in this study was 0.41 (SD=0.25). This is remarkably similar to the TRANSFORM-2 baseline value (0.417), despite having a lower average MADRs total score of 32. As noted above, this is close to the mean MADRS score of patients who relapsed in SUSTAIN-1.

3. Previous NICE models have relied on even more significant differences between depressive and remission health states than we are using. The NICE CG90 model used utilities of 0.33 for moderate depression, 0.15 for severe depression, 0.85 for remission and 0.72 for response (47). The model outcomes were used by NICE to inform CG90 recommendations. This is a greater difference to the current utility values, which are 0.43 for MDE, 0.86 for remission, and 0.77 for response.

4. Furthermore, the utility for MDE remains constant throughout all treatment phases in the model, despite additional non-response to subsequent treatments. This is conservative, as after more failed treatments (spending more time in the non-response state) the quality of life for patients often deteriorates independently of the level of depressive symptoms (1). Literature shows this is partly due to the emotional burden associated with inadequate response and the impact of trying numerous medications has on patients. The longer the current episode duration, the more likely the respondents experienced feelings of frustration with medication (48, 49). By assuming MDE utility remains constant when patients do not respond to subsequent treatments, the approach taken in the economic model is conservative.

The information provided above aims to address the Committee's concern, and we ask the Committee to reconsider their conclusions on whether the baseline utility from TRANSFORM-2 is appropriate to inform the MDE health state. If the above does not sufficiently address the Committee's concern, we have provided an additional scenario below.

3.1: Amending the criteria for MDE health state allows consideration of an alternative utility from a QoL study conducted in UK patients with TRD

Given the Committee's concerns with the use of the baseline utility from TRANSFORM-2, and differences in thresholds used for the MDE health state in the model and the clinical trials, we propose a scenario where the criteria for the MDE health state is amended. This allows consideration from the results from a cross-sectional study of UK patients with TRD who cover the full moderate-severe depressive spectrum. Patients with TRD in the study had a PHQ-9 score of ≥ 13.8 (corresponding to a MADRS total score of ≥20) and had a mean utility of 0.430. This change in criteria of the MDE health state covers the relapse threshold (≥22) used in SUSTAIN-1, and as such mitigates the concerns with the different thresholds used in the clinical trial programme and model. See Appendix D for further details on this study and Section 5.0 for revised ICERs including this data source for the MDE utility.

Overall, relapse signifies the transition from remission to a MDE health state. Other studies have shown the utility to inform the MDE health state (even when considering different thresholds for the MDE health state) to be reasonable and, as such, the difference between health states not to be overestimated. Available evidence suggests that as ineffective treatment lines increase, the quality of life of patients decreases further (1, 50) and therefore this could be considered a conservative assumption.

4.0: It is appropriate to assume different healthcare resource use costs per health state, which could lead to different medical costs between treatment arms. We propose a sensitivity analysis with reduced cost differences among health states to address the Committee's concerns and include additional costs that may be associated with commissioning ESK-NS

Key Point 4 summary

- It is not appropriate to use SUSTAIN-1 to inform HCRU per treatment arm as it was not designed to collect resource use data, cannot provide any conclusions due to a small number of events, does not consider the full cohort of patients with TRD and is not generalisable to resource use in UK clinical practice. Evidence shows that differential costs per health state are appropriate. This has also been the approach used in previous NICE decision making.
- Given the Committee's concerns with the MDE health state, we provide a sensitivity analysis where the costs of the health states are adjusted using the lower bound of the 95% CI costs, resulting in reduced cost differences among health states.
- A sensitivity analysis, with revised health state costs, is provided to address the Committee's concerns, such that it is acceptable to have differential costs per health state.
- Estimates of the costs of commissioning are incorporated into the model and have very minimal impact on the cost effectiveness.

ACD (Section 3.28): 'The committee considered that these costs were driven by events and that there is considerable uncertainty whether esketamine would reduce these events from the SUSTAIN-1 data. Because of this, and the importance of the overpopulated MDE health state, it concluded that it was most appropriate to make healthcare resource use costs equal across treatment arms. The committee did not consider this conservative because resource use of esketamine could be higher than placebo if using SUSTAIN-1 data, and there is considerable uncertainty with this assumption.'

Janssen response:

In the previous sections 2 and 3, we have addressed the Committee's concerns regarding the potential overpopulation of the MDE health state. Section 2 outlines how the proportion of patients in the MDE health state reflects the long-term outcomes of patients with TRD when considering evidence from the wider literature. Section 3 explains the appropriateness of the MDE health state, and specifically the MDE health state utility. Given these considerations, the scenario where equal healthcare resource use is assumed is not logical or clinically appropriate to use and further explanation is provided below.

4.1. It is not appropriate to use SUSTAIN-1 to inform HCRU per treatment arm and evidence provided shows differential costs per health state is appropriate

The ACD states that the rationale for the equal healthcare costs is based on the SUSTAIN-1 relapse data, but this is inappropriate for several reasons. The SUSTAIN-1 data show that the primary reason for relapse in stable remitters was worsening depression manifested as a deteriorating MADRS total score, with few patients meeting criteria for relapse based on a clinically relevant event or hospitalisation. As noted by the Committee, hospitalisation as the reason for relapse event in the ESK-NS study arm was reported in three patients (n=3) versus no hospitalisation in the OAD+PBO-NS arm (n=0) for patients who were in stable remission. This is then used as part of the rationale for modelling equal medical costs between arms. This rationale for using this to model equal medical costs between arms is flawed for a number of reasons; SUSTAIN-1 was not designed to collect resource use data, conclusions cannot be

made due to a small number of events, the full cohort of patients with TRD is not considered and SUSTAIN-1 resource use is not generalisable to the UK. This is outlined below:

- SUSTAIN-1, as a relapse prevention trial, was designed to assess relapse events as soon as they occurred. Observation of healthcare resource utilisation (HCRU) of patients once they relapsed was only very limited as patients were followed up for only up to two weeks post relapse. For stable remitters, the mean follow-up durations are 143 days pre-relapse vs 13 days post-relapse. As such, potential subsequent hospitalisations and other HCRU for the relapsed patients were not captured after end of follow-up in SUSTAIN-1.
- Consequently, there are small number of hospitalisation events (n=3 and n=1) due to the insufficient duration of observation and follow up in SUSTAIN-1. This is likely to be a chance finding. This is acknowledged in the ACD, as it states that SUSTAIN-1 was not appropriate to detect differences in healthcare resource events.
- In contrast to the stable remitters data set, the HCRU data from the SUSTAIN-1 stable responders set does not support this assumption. There were 0 hospitalisations in the ESK-NS + OAD arm compared to 1 hospitalisation in the OAD + PBO-NS arm. Similar as for the stable remitters, the number of hospitalisations are too small to make any inference for the HCRU of the study arms.
- SUSTAIN-1 was an international multi-centre trial and so the HCRU data collected is not generalisable to the UK. This is acknowledged in the ACD.
- The model includes the full cohort of patients. In contrast, given the study design and objective, the study included only stable responders and stable remitters after 16 weeks of ESK-NS treatment. Patients not included in the SUSTAIN-1 trial (non-responders) are likely to have much higher HCRU than remitters. Patients on the OAD had fewer remitters and responders after induction treatment, so would be expected to have higher HCRU in clinical practice.
- The broader implication of using this assumption is not aligned to clinical evidence or feedback from clinical experts. The assumption of equal non-drug medical costs means that patients who are symptomatic and in an active depressive episode have the same HCRU and costs (including hospitalisation) as patients with TRD who are in remission or recovery is not considered clinically reasonable and lacks clinical validity.

For the reasons outlined, we believe that the SUSTAIN-1 HCRU data is inappropriate to inform the costs per health state in the model. A more appropriate approach that is aligned to previous depression models, reflective of the data in TRD and clinically plausible is to apply HCRU costs by health state, as done in the company base analysis. Further details are explained in this section below.

In the company base case analysis, a robust retrospective chart review study was used that found that HCRU and costs for UK patients with TRD vary substantially by health state. This is reflective of the available literature and validated by clinical opinion that patients in remission and recovery have a lower HCRU and hence medical cost associated with their disease management compared with patients in an MDE state. This study was comprehensive and included data from 295 patients with TRD from regions across the UK, ensuring a representative real world patient population. To enhance the generalisability, physicians contributing to the study were recruited from a diverse geographical spread and mixed primary/secondary care practice. The chart review captured the HCRU and estimated the costs of patients with TRD, including the following resources:

- Consultations in primary and secondary care (e.g., GPs and psychiatrists),
- Use of Crisis Resolution Home Treatment Teams (CRHTT),
- Use of non-drug treatments, such as counselling or psychotherapy,

- Any hospitalisations, including time spent in ICU or on a psychiatric ward.

The data from this study are used to inform the costs per health state in the base case. Based on the systematic review used to support the company submission, this study is the only study to inform how HCRU associated with TRD in the UK varies across different health states. It is therefore the most generalisable evidence and methodologically robust evidence to inform Committee decision making. The findings from a similar study conducted in Belgium had the same conclusion (51) and therefore show the study to have face validity.

In addition to the UK retrospective chart review study, consistent findings from the literature further characterises the association between the increased burden of TRD and elevations in both direct and indirect costs. A systematic review conducted in 2019 (1) identified a consistent trend with increasing medical costs and decreasing HRQoL when the severity of TRD and non-response rate increases within an MDD episode. This is consistent both the 2009 review conducted by Mauskopf et al. (51) and the 2014 review by Mrazek et al (50).

The assumption of different costs per health state has been used consistently in other depression models used by NICE. Although not reviewing a TRD population, the Committee in TA367 accepted the same approach as per the company base case (differential costs per health state), reflecting the decrease in resource use of patients who respond/remit to treatment. The economic model used to support decision making in NICE CG90 of the cost effectiveness of pharmacological interventions implemented different costs based on response to treatment too. NICE have therefore consistently made their decisions previously on the basis of costs differ by health state in modelling of depression interventions.

The evidence provided above shows that the approach to use differential costs per health state is more appropriate than assuming equal medical costs in each treatment arm. We therefore request the Committee to re-consider on the basis of the clinical plausibility, evidence from the literature and precedence and consistency of models that NICE have considered in this disease area previously to use medical resource use costs that are linked to the patient's health state.

4.2 Sensitivity analysis with reduced cost differentials between health states

The use of equal medical costs is not appropriate in the economic model for a number of reasons, as noted above. We recognise, however, that the Committee were not certain with the cost estimates from the retrospective chart review. We feel that considering the potential positioning of ESK-NS later in the pathway (as per Section 6.1) would mean the results of retrospective chart review would be even conservative, as patients later in the pathway are associated with greater HCRU and medical costs (1). Nevertheless, we provide a conservative sensitivity analysis using a reduced differential cost per health state to address the Committee's concerns regarding some of the costs included in the retrospective chart review. A scenario using the lower 95% CI for all health states would utilise mean costs of £761.48 (MDE), £102.81 (remission and response), £47.97 (recovery) compared to mean costs of £980.08, £164.46, and £83.75 respectively (Table 5).

Table 5: Base case and 95% CI health state costs

	Base case health state costs	Lower 95% CI health state costs
MDE	£980.08	£761.48
Remission	£164.46	£102.81
Response	£164.46	£102.81
Recovery	£83.75	£47.97

When using these inputs, the ICER for ESK-NS is [redacted] (including carer disutility) to [redacted] (excluding carer disutility). This is provided as a conservative scenario analysis to address the Committee's concerns, such that it is acceptable to have differential costs per health state. The implication of using the lower bound of the 95% CI value for the health state costs would be to ignore the patients who require the most HCRU, which would still be treated in the real-world NHS. As such, we believe the scenario with the base case health state costs to be most appropriate, as they are based on the mean observed data.

4.3: Estimates of the costs of commissioning are incorporated into the model and have very minimal impact on the cost effectiveness

ACD, Section 3.30: *'Significant investment will be needed to use esketamine in the NHS, but costs are difficult to quantify'*

ACD, Section 3.31: *'It will take time and resource use for esketamine to become part of clinical practice'*

The commissioning expert stated that most mental health services are not well established to offer ESK-NS administration and post-dose monitoring. Adoption of the use of ESK-NS will require adjustments in the configuration of services for people with TRD. The expert also stated that the following investments need to be considered to introduce the technology:

- Costs of conversion of ECT suites,
- Costs of medical equipment to monitor and manage any post-dose medical complications
- staff training to manage post-dose complications, including potential costs of recruitment if there are not enough staff currently available in practice,
- Costs associated with the controlled nature of the drug, including storage, transportation, disposal and adequate staffing and governance training,
- Costs associated with creating and managing a registry to avoid misuse and abuse of esketamine.

A survey of 16 Mental Health Trusts (MHTs) was conducted. Full results for the survey are available in Appendix E. The majority of MHTs did not identify any additional costs of commissioning. A summary of the survey for each of the cost element is provided below.

Costs of conversion of ECT suites

The majority of MHTs are planning on utilising existing clinics or ECT suites, without any additional infrastructure costs. One MHT that is looking at renewing infrastructure is for the development of a whole new service for rTMS, which could also include ESK-NS. These costs, however, should not be included in any additional costs of commissioning for ESK-NS as they would occur anyway.

Costs of medical equipment

These clinics already have the medical equipment required, so they see no additional cost of introducing an ESK-NS service. If the service was expanded to community settings over a long period of time then additional costs of medical equipment may be required.

Costs associated with the controlled nature of the drug
A few areas of additional cost were identified, but mainly for a larger or additional controlled drug (CD) cabinets.

Costs associated with staff training
Those that will repurpose the existing clinic said they will not necessarily require additional nurse support, as this is not incremental to existing resources. It was highlighted that in the first few years after a recommendation from NICE, the numbers of patients treated with ESK-NS will be limited, therefore setting up new clinics (with new costs) is not required.

Costs associated with creating and managing a registry to avoid misuse and abuse of esketamine.

See section 1.7 for details on the proposed registry.

The total costs for all 16 MHTs included in the survey was £2100 for CD cabinets. If this sample is extrapolated to cover all 69 MHTs and split by the total expected number of patients who will be treated with ESK-NS in the first 5 years, it would result in a cost of £1.62 per patient, which has now been included in the model scenarios reported in this document. New infrastructure or additional costs could be introduced after many years, if the numbers of patients increase, as MHTs may increase clinic capacity. These additional costs, however, cannot be costed at present because this is an unknown and it is uncertain.

5.0: Revised post ACD- 2 scenarios for consideration (Full population)

In this section, revised scenarios are presented for the Committee's consideration. We have included the following Committee preferred assumptions from the ACD in this scenario:

- Use of unadjusted short-term data
- 20-year time horizon
- 2:1 administration cost
- Removed mortality effect
- Committee preferred treatment discontinuation
- ERG method of including carer disutility

The base case model assumptions result in conservative long term model outcomes compared to the observed STAR*D study and other cohort studies, as noted in Section 2.2. Given the Committee's concerns, however, a revised method for subsequent treatments is used in the scenario below, which is aligned to the ERG's revised approach (see Section 2.1). This approach does not amplify the high efficacy observed in the TRANSFORM-2 OAD-PBO-NS arm as with the original ERG assumptions that were used in the second ACD preferred Committee assumptions.

In addition to this revised method for subsequent treatments, we maintain the assumptions of differential medical costs per treatment arm by using differential costs per health state and per treatment arm. A reduction in the differences between health states are provided as a sensitivity analysis using the lower bound of the 95% CI values from the retrospective chart review, which shows ESK-NS remains cost effective when using this conservative assumption.

Given the Committee's concerns with the different thresholds used in the model and clinical trial programme, we have explored amending the criteria of the MDE health state. This allows consideration from the results from a cross sectional study of UK quality of life patients with TRD. The revised ACD scenario incorporates the utility from this study for the MDE health state.

We note that NICE has included a reduced treatment discontinuation in the revised NICE ACD base case, with 0% non-efficacy discontinuation after 9 months in remission and 70% discontinuation by 2 years, which are included in the scenarios below. This is conservative given the number of patients to stay on treatment beyond two years is approximately twice as high (30% vs 16%) as estimated by 25 UK psychiatrists. We suggest it may be more appropriate to use the company base case treatment discontinuation for decision making. The below scenarios (Table 6), however, include the NICE ACD base case assumptions on treatment discontinuation.

Table 6: Revised post ACD- 2 scenarios for consideration (Full population)

Key differential parameters	NICE ACD 2 preferred scenarios	Janssen revised ACD scenario	Janssen revised ACD 2 scenario: sensitivity analysis
MDE utility	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRD QoL study (MADRS total score of ≥20)
Medical (HCRU) costs	Equal medical costs per treatment arm	Differential medical costs per treatment arm based on health state costs	95% CI lower bound of health state costs
Subsequent treatment approach	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy
Best Supportive care approach	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy
Additional costs of commissioning	Not included	Included (see Section 4.3)	Included (see Section 4.3)
ICER	[redacted]	[redacted]	[redacted]

*= excluding carer disutility

The specific parameters to inform the revised scenario are provided in Appendix F. Further scenarios exploring the impact of different ESK-NS dosing on the ICER, which uses the 95% CI for the dosing from the trials, are provided in Appendix H.

Overall, the ICER for ESK-NS in the full licensed population reduces to [redacted](including carer disutility) [redacted](excluding carer disutility) when these changes are applied. We consider this should provide the basis of the revised base case ICER for the Committee to consider.

5.1 Revised scenario at later line positioning: non-response to at least 3+ prior OAD.

ACD Section 3.4: *'Esketamine is likely to be used later in the treatment pathway because it has a higher treatment burden than other treatments'*

All of the data and scenarios presented so far in the NICE process comprise of the full label population (TRANSFORM-2 population), which includes patients who have failed at least two prior OADs.

We note that multiple clinical consultees and the clinical expert have stated that ESK-NS is likely to be used later in the treatment pathway, as stated in the ACD. Specifically, clinical experts say that ESK-NS will be used in patients who have failed at least three prior OADs, given that the current treatment pathway of depression in the NHS includes late referrals to a secondary care setting. For patients who have failed at least three prior treatments, there is a higher unmet need given the reduced treatment options and lower likelihood of responding to each successive treatment (5). As shown in the literature, the quality of life and burden of disease increases significantly as patients do not respond to additional treatments (1, 50).

After non-response to 3 prior treatments, it can be expected that augmentation with antipsychotics or lithium would become more relevant comparators for ESK-NS, besides monotherapy oral antidepressants. Published meta-analyses have shown the beneficial effect size of ESK-NS compared to augmented antidepressant therapies. The effect size was nearly twice as high versus antidepressant augmentation with second-generation antipsychotics (17). The augmentation of atypical antipsychotics may be effective as adjunctive therapy; however, their adverse effect profile may be unfavourable to some patients (53).

Given the most recent ACD and the feedback from clinical experts as a result of the 1st NICE ACD consultation, we have conducted an analysis of the clinical efficacy and cost effectiveness, when considering the subgroup of patients who have not responded to at least three oral antidepressants in the clinical trial programme. The primary outcome of the TRANSFORM-2 study (change from baseline MADRS to Day 28) was a prespecified subgroup, as shown in Appendix E.1.1 of the Company Submission. The results from the subgroup analysis, including response and remission, can be found below.

Clinical efficacy: Non-response to at least 3+ prior treatments

Short term efficacy: TRANSFORM-2

The change from baseline MADRS to day 28 for the subgroups of those who did not respond to at least 3 prior OAD treatments is shown below for TRANSFORM-2 (Table 7). [redacted] in the ESK-NS + OAD arm and [redacted] patients in the OAD + PBO-NS arm had non-response to 3 or more prior OAD treatments within the current episode.

The results from this subgroup show that the difference between study arms at day 28 was [redacted] in favour of ESK-NS + OAD, [redacted]

Table 7: Change from baseline MADRS

	ESK-NS + OAD		OAD + PBO-NS		Difference in LS Mean CFB	
	N	LS Mean CFB(SE) [95%CI], p-value	N	LS Mean CFB(SE) [95%CI], p-value	Estimate[95%CI], p-value	p-value (Interaction)
All patients	101	-19.8 (1.250); [-22.3, -17.3], <.0001	100	-15.8 (1.260); [-18.3,-13.3], <.0001	-4.0 [-7.31, -.636], 0.0199	NE
Non-response to 2 prior treatments	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response at least 3 prior treatments	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	

The remission and response outcomes for this subgroup are shown below in Table 8.

Table 8: TRANSFORM-2 remission and response per 3+ prior failures subgroup

	Remission			Response		
TRANSFORM-2 remission	ESK-NS + OAD remission	OAD + PBO-NS remission	p-interaction	ESK-NS + OAD response	OAD + PBO-NS response	p-interaction
All patients	52.5% (n=53/101)	31.0% (n=31/100)		69.3% (n=70/101)	52.0% (n=52/100)	NE
Non-response to 2 prior OADs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response to at least 3 prior OADs	[redacted]	[redacted]		[redacted]	[redacted]	

[redacted]

[redacted]

[redacted]

Figure 3: TRANSFORM-2 mean change in MADRS from baseline to Day 28 for patients who did not respond to at least 3 prior OAD treatments

[redacted]

[redacted] This is supportive of the differential mechanism of action for ESK-NS compared to existing antidepressants. This plausibility is acknowledged in previous feedback to NICE in TA367:

TA367, FAD, p42: *'However, the clinical expert noted that the relative effectiveness of the antidepressants compared with one another may also change at each subsequent line of treatment. The clinical expert explained that depression which does not respond to 1 or 2 SSRIs may be mediated by different receptors, so the relative effectiveness of treatments with a different mechanism may differ across subsequent lines of treatment.'*

<u>Cost effectiveness: Non-response to at least 3+ prior treatments</u>

<u>Overall response and remission (pooled and weighted TRANSFORM-2 and TRANSFORM-3) for the subgroup of patients who did not respond to at least 3 prior treatments</u>

The overall response and remission for the model scenario in this subgroup is shown below in Table 9. This includes the weighted use of the TRANSFORM-2 and TRANSFORM-3 data.

Table 9: Overall response and remission for the subgroup of patients who did not respond to at least 3 prior treatments

	Remission		Response	
	ESK-NS + OAD remission	**OAD + PBO-NS remission**	**ESK-NS + OAD response**	**OAD + PBO-NS response**
All patients	46.1%	26.6%	15.5%	18.4%
Non-response to at least 3 prior OADs	[redacted]	[redacted]	[redacted]	[redacted]

The clinical data for the TRANSFORM-3 subgroup of those who did not respond to at least 3 prior OAD treatments is shown in Appendix G, as well as other inputs to inform this scenario.

[redacted] When inputting the data in the model, the ICER for this scenario can be found below in Table 10. Overall, the [redacted] (Including carer disutility) to xxxxxxxxx (excluding carer disutility) when the positioning of ESK-NS is considered one treatment failure later in the pathway. This ICER would increase to xxxxxxxxx (including carer disutility) to xxxxx (excluding carer disutility) if the alternative MDE utility (0.43) and the 95% CI lower bound health state costs (see section 3.1 and 4.2 for further details) are used.

Table 10: Revised scenarios at later line positioning: non-response to at least 3+ prior OAD.

Key differential parameters	**Janssen revised ACD scenario**	**Janssen revised ACD scenario at later line positioning (non-response to at least 3 prior OADs)**	**Later line positioning (non-response to at least 3 prior OADs) with different utility and health state costs**
MDE utility	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRD QoL study (MADRS total score of ≥20)
Medical (HCRU) costs	Differential medical costs per treatment arm based on health state costs	Base case medical cost per health state	95% CI lower bound of health state costs
Subsequent treatment approach	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy
Best Supportive care approach	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy
Additional costs of commissioning	Included (see Section 4.3)	Included (see Section 4.3)	Included (see Section 4.3)
ICER	[redacted]	[redacted]	[redacted]

6.0: Other points and Errors/ factual inaccuracies

Table 11 below presents additional points and errors.

Table 11: Other points and factual inaccuracies

Location in ACD and statement	Rationale
Section 3.12 *'the company also did not use data from SUSTAIN-1 for relapse rate in the oral antidepressant with placebo arm in the economic model to avoid any withdrawal effect.'*	STAR*D was used to inform the relapse rate in the OAD arm to avoid the potential bias that could occur when using the SUSTAIN-1 OAD + PBO-NS data due to its ESK-NS withdrawal study design. However, as an alternative scenario, data from SUSTAIN-1 has been used to provide alternative estimates or relapse and loss of response for OAD (Section B.3.4.4.8 of company submission). This decreases the ICER compared to the use of the STAR*D data.
Section 3.5 *'The ERG added that the network meta-analysis only used adjusted effects for the oral antidepressant with placebo arm of esketamine.'*	The network meta-analysis using the unadjusted effects were provided to NICE in July 2019 in Section D.1.3.4 of the appendices and the Company response to Question C2 of the ERG Clarification Questions.
Section 3.6 *'An expert from the NICE guideline on depression noted that psychological therapies were not included as comparators or with combination treatments in the company's submission but were included in the NICE appraisal scope.'*	This is an error, as psychological therapies were not included as a comparator in the NICE appraisal scope.
Section 3.8: *'The committee also noted that the score used for relapse was not equivalent to the MADRS score for moderate to severe depression, which affected the health state utility values and transitions in the economic model.'*	A score of 20-34 on the MADRS scale is regarded as moderate severity. The threshold used for relapse, a MADRS score of 22 or higher, is consistent with presence of MDE symptoms of moderate severity.
Section 3.15: *'The committee concluded that it had not seen evidence that the additional clinical contact involved in the placebo arm improved clinical outcomes.'*	This is not aligned to the evidence provided, which shows that the improvement in clinical outcomes from clinical contact does not rely on CBT to improve clinical outcomes. This was demonstrated in the Posternak study (54).
Section 3.17: *'The model output suggests that within 1 year, 78% of people with treatment-resistant depression in current clinical practice do not have symptom response to any treatments long-term. So, they then occupy the MDE state for the remainder of the time horizon.'*	The statement regarding patients who occupy the MDE health state for the remainder of the time horizon is incorrect, as patients who are in the BSC treatment phase have an ongoing likelihood of achieving response (3.9% chance per 4 weeks) or remission (10.5% per 4 weeks) and hence a proportion of patients will continuously move out of the MDE state.
Section 3.21: *'The transitions between response and remission states were also sourced from STAR*D for both arms, although this assumption was not fully explored by the company.'*	This is a factual inaccuracy. Data from SUSTAIN-1 were used to inform the rate of transition from response to remission, as noted in Section B.3.2.9.2.1 of the company submission.

Section 3.21: '*The relapse and loss of response rates for the oral antidepressant arm were sourced from the STAR*D trial. The STAR*D trial used different relapse criteria. Also, it was unclear if the population from STAR*D is generalisable to the NHS.'*	Relapse is defined as a return of the MDE following the achievement of remission but before fulfilling the criteria for recovery from the current episode. The definitions of relapse used in STAR*D and SUSTAIN-1 are appropriate to capture this worsening of depressive symptoms. In STAR*D, relapse was declared when the QIDS-SR16 score collected by the interactive voice response system during the follow up phase was ≥11 (corresponding to an HRSD17 ≥14). This definition correlates to moderate to severe depression, which is similar to the criteria in the SUSTAIN-1 trial.
Section 3.23: *'The committee noted that the mean EQ-5D-5L health score index was consistently higher than 0.8 for all participants at the end of maintenance for SUSTAIN-1. In participants who were randomised to withdraw from esketamine, 45% of people whose depression was in stable remission and 58% of people whose depression was in stable response relapsed. The committee considered that this would not correspond to the relatively high EQ-5D-5L health score index above 0.8 if this represented a true transition to the MDE health state.'*	The SUSTAIN-1 utility data for stable responders and stable remitters do not capture patients who transitioned to the MDE health state. SUSTAIN-1 only included stable remitters and stable responders. If patients relapsed or lost response, then they no longer contributed to these data sets. This can be seen as the sample size in the SUSTAIN-1 stable remitters and stable responders reduces over time. Patients who are in stable remission or stable response can be expected to have consistently higher utility scores, and the SUSTAIN-1 study data are consistent with this.
Section 3.27 '*The committee would like to see the proportion of people having each dose, how often people have esketamine (weekly or every 2 weeks), reasons for the dosing choices.'*	Please see Appendix H for further scenarios exploring the impact of different dosing scenarios. The reasons for the dosing choices were previously provided to the ERG (see response to ERG clarification questions, Section A.9).
Section 3.28 *'The committee considered that CBT and ECT were excluded from the trials and should not be included in the medical costs.'*	CBT and ECT should not be excluded from the costs of the health states. Whilst the retrospective chart review showed that CBT and ECT do not comprise a large proportion of the costs of patients with TRD, it is not appropriate to exclude these costs, which are still incurred in the NHS.
Section 3.30 *'costs associated with creating and managing a registry to avoid misuse and abuse of esketamine.'*	Costs to the NHS will be minimal as the company has agreed with the MHRA to cover the costs of the data collection. Nurse time for the administration of the registry will be captured during supervision of the patient.
Section 3.31 *'They said a reasonable time to implement esketamine in a community setting would be 12 months, and 6 months in a secondary hospital clinic setting.'*	Feedback from multiple mental health trusts indicates 180 days is not required. Feedback from NHS at a Trust level has clearly said that significant infrastructure investments are not required. 82% of the sites said that they will repurpose existing premises for the adoption of ESK-NS into the NHS.

References

1. Johnston, K., Powell, L., Anderson, I., Szabo, S., & Cline, S. (2019). The burden of treatment-resistant depression: A systematic review of the economic and quality of life literature. Journal Of Affective Disorders, 242, 195-210. doi: 10.1016/j.jad.2018.06.045
2. Briley, M., & Lépine. (2011). The increasing burden of depression. Neuropsychiatric Disease And Treatment, 3. doi: 10.2147/ndt.s19617
3. Gupta, M., Holshausen, K., Best, M., Jokic, R., Milev, R., & Bernard, T. et al. (2013). Relationships Among Neurocognition, Symptoms, and Functioning in Treatment-Resistant Depression. Archives Of Clinical Neuropsychology, 28(3), 272-281. doi: 10.1093/arclin/act002
4. Kings Fund. 2020. Has the government put mental health on an equal footing with physical health?. Retrieved 16 October 2020, from https://www.kingsfund.org.uk/projects/verdict/has-government-put-mental-health-equal-footing-physical-health
5. Rush, A., Trivedi, M., Wisniewski, S., Nierenberg, A., Stewart, J., & Warden, D. et al. (2006). Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. American Journal Of Psychiatry, 163(11), 1905-1917. doi: 10.1176/ajp.2006.163.11.1905
6. Wong, C., Siah, K., & Lo, A. (2018). Estimation of clinical trial success rates and related parameters. Biostatistics, 20(2), 273-286. doi: 10.1093/biostatistics/kxx069
7. Khin NA, Chen YF, Yang Y, Yang P, Laughren TP. Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. J Clin Psychiatry 2011; 72: 464-72.
8. Kesselheim, A., Hwang, T., & Franklin, J. (2015). Two decades of new drug development for central nervous system disorders. Nature Reviews Drug Discovery, 14(12), 815-816. doi: 10.1038/nrd4793
9. Hieronymus, F., Jauhar, S., Østergaard, S., & Young, A. (2020). One (effect) size does not fit at all: Interpreting clinical significance and effect sizes in depression treatment trials. Journal Of Psychopharmacology, 34(10), 1074-1078. doi: 10.1177/0269881120922950
10. Duru G, Fantino B. The clinical relevance of changes in the Montgomery-Asberg Depression Rating Scale using the minimum clinically important difference approach. Curr Med Res Opin. 2008;24(5):1329-3
11. Liebowitz, M., Lam, R., Lepola, U., Datto, C., Sweitzer, D., & Eriksson, H. (2010). Efficacy and tolerability of extended release quetiapine fumarate monotherapy as maintenance treatment of major depressive disorder: a randomized, placebo-controlled trial. Depression And Anxiety, 27(10), 964-976. doi: 10.1002/da.20740
12. Montgomery, S., & Möller, H. (2009). Is the significant superiority of escitalopram compared with other antidepressants clinically relevant?. International Clinical Psychopharmacology, 24(3), 111-118. doi: 10.1097/yic.0b013e32832a8eb2
13. Melander, H., Salmonson, T., Abadie, E., & van Zwieten-Boot, B. (2008). A regulatory Apologia — A review of placebo-controlled studies in regulatory submissions of new-generation antidepressants. European Neuropsychopharmacology, 18(9), 623-627. doi: 10.1016/j.euroneuro.2008.06.003
14. Singh, J., Daly, E., Mathews, M., Fedgchin, M., Popova, V., Hough, D., & Drevets, W. (2020). Approval of esketamine for treatment-resistant depression. The Lancet Psychiatry, 7(3), 232-235. doi: 10.1016/s2215-0366(19)30533-4
15. CHMP. Spravato Assessment Report. Available from : https://www.ema.europa.eu/en/documents/assessment-report/spravato-epar-public-assessment-report_en.pdf
16. Response to Moncrieff & Horowitz: Evidence Based Medicine or A Brave New World?. (2020). Retrieved 18 September 2020, from https://www.bmj.com/content/366/bmj.l5572/rr-5
17. Dold, M., Bartova, L., & Kasper, S. (2020). Treatment Response of Add-On Esketamine Nasal Spray in Resistant Major Depression in Relation to Add-On Second-Generation Antipsychotic Treatment. International Journal Of Neuropsychopharmacology, 23(7), 440-445. doi: 10.1093/ijnp/pyaa034
18. G.Katz, E., Hough, D., Doherty, T., Lane, R., Singh, J., & Levitan, B. (2020). Benefit-Risk Assessment of Esketamine Nasal Spray versus Placebo in Treatment-Resistant Depression. Clinical Pharmacology & Therapeutics. doi: 10.1002/cpt.2024
19. Citrome, L., DiBernardo, A., & Singh, J. (2020). Appraising esketamine nasal spray for the management of treatment-resistant depression in adults: Number needed to treat, number needed to harm, and likelihood to be helped or harmed. Journal Of Affective Disorders, 271, 228-238. doi: 10.1016/j.jad.2020.03.106

20. Zheng, W., Cai, D., Xiang, Y., Zheng, W., Jiang, W., & Sim, K. et al. (2020). Adjunctive intranasal esketamine for major depressive disorder: A systematic review of randomized double-blind controlled-placebo studies. Journal Of Affective Disorders, 265, 63-70. doi: 10.1016/j.jad.2020.01.002
21. Kim, J., Farchione, T., Potter, A., Chen, Q., & Temple, R. (2019). Esketamine for Treatment-Resistant Depression — First FDA-Approved Antidepressant in a New Class. New England Journal Of Medicine, 381(1), 1-4. doi: 10.1056/nejmp1903305
22. Papakostas, G., Salloum, N., Hock, R., Jha, M., Murrough, J., & Mathew, S. et al. (2020). Efficacy of Esketamine Augmentation in Major Depressive Disorder. The Journal Of Clinical Psychiatry, 81(4). doi: 10.4088/jcp.19r12889
23. European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment of depression. EMA/CHMP/185423/2010 Rev. 2. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-depression_en.pdf.
24. Cleare, A., Pariante, C., Young, A., Anderson, I., Christmas, D., & Cowen, P. et al. (2015). Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2008 British Association for Psychopharmacology guidelines. Journal Of Psychopharmacology, 29(5), 459-525. doi: 10.1177/0269881115581093
25. Taylor, D. Barnes, T., Young, A. The Maudsley Prescribing Guidelines in Psychiatry. 13th Edition.
26. Khin NA. Update on FDA exploratory analyses of aggregated efficacy data from antidepressant trials. Presented at NCDEU Regulatory Plenary. New FDA & EMA initiatives in Alzheimer's disease, depression and schizophrenia. May 30, 2013
27. Yang P. Recent advances in design and methodology in psychiatric clinical trials. Presented at: Joint Statistical Meetings, Montréal, QC Canada, August 3-8, 2013
28. US Department of Health and Human Services, Food and Drug Administration. Major Depressive Disorder: Developing Drugs for Treatment. Guidance for Industry. Draft June 2018. Available at https://www.fda.gov/media/113988/download.
29. Yieh L, Aluisio L, Wajs E, et al. Withdrawal symptom assessment – esketamine nasal spray: open-label safety study in treatment-resistant depression. Presented at: Psych Congress 2019; October 3-6, 2019; San Diego, CA. Poster 242.
30. Sackeim HA, Haskett RF, Mulsant BH, et al. Continuation pharmacotherapy in the prevention of relapse following electroconvulsive therapy: a randomized controlled trial. JAMA. 2001;285(10): 1299-1307.
31. Sackeim, H., Prudic, J., Devanand, D., Decina, P., Kerr, B., & Malitz, S. (1990). The Impact of Medication Resistance and Continuation Pharmacotherapy on Relapse following Response to Electroconvulsive Therapy in Major Depression. Journal Of Clinical Psychopharmacology, 10(2), 96-104. doi: 10.1097/00004714-199004000-00004
32. Bernal, M., Haro, J., Bernert, S., Brugha, T., de Graaf, R., & Bruffaerts, R. et al. (2020). Risk factors for suicidality in Europe: Results from the ESEMED study. Journal Of Affective Disorders, 101(1-3), 27-34. Retrieved from https://www.academia.edu/10048362/Risk_factors_for_suicidality_in_Europe_Results_from_the_ESEMED_study
33. Dunner, D., Rush, A., Russell, J., Burke, M., Woodard, S., Wingard, P., & Allen, J. (2006). Prospective, Long-Term, Multicenter Study of the Naturalistic Outcomes of Patients With Treatment-Resistant Depression. The Journal Of Clinical Psychiatry, 67(05), 688-695. doi: 10.4088/jcp.v67n0501
34. Conway, C., Kumar, A., Xiong, W., Bunker, M., Aaronson, S., & Rush, A. (2018). Chronic Vagus Nerve Stimulation Significantly Improves Quality of Life in Treatment-Resistant Major Depression. The Journal Of Clinical Psychiatry, 79(5). doi: 10.4088/jcp.18m12178
35. Conway, C., Kumar, A., Xiong, W., Bunker, M., Aaronson, S., & Rush, A. (2018). Chronic Vagus Nerve Stimulation Significantly Improves Quality of Life in Treatment-Resistant Major Depression. The Journal Of Clinical Psychiatry, 79(5). doi: 10.4088/jcp.18m12178
36. Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. J Clin Psychiatry. 1997; 58(Suppl 13):23–29.
37. Rush, A., Aaronson, S., & Demyttenaere, K. (2018). Difficult-to-treat depression: A clinical and research roadmap for when remission is elusive. Australian & New Zealand Journal Of Psychiatry, 53(2), 109-118. doi: 10.1177/0004867418808585
38. Sackeim, H. (2016). Acute Continuation and Maintenance Treatment of Major Depressive Episodes With Transcranial Magnetic Stimulation. Brain Stimulation, 9(3), 313-319. doi: 10.1016/j.brs.2016.03.006
39. Fonagy, P., Rost, F., Carlyle, J., McPherson, S., Thomas, R., & Pasco Fearon, R. et al. (2015). Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock Adult Depression Study (TADS). World Psychiatry, 14(3), 312-321. doi: 10.1002/wps.20267

40. Aaronson, S., Sears, P., Ruvuna, F., Bunker, M., Conway, C., & Dougherty, D. et al. (2017). A 5-Year Observational Study of Patients With Treatment-Resistant Depression Treated With Vagus Nerve Stimulation or Treatment as Usual: Comparison of Response, Remission, and Suicidality. American Journal Of Psychiatry, 174(7), 640-648. doi: 10.1176/appi.ajp.2017.16010034
41. Janssen. Data on File. October 2020.
42. Edwards SJ, Hamilton V, Nherera L, Trevor N. Lithium or an atypical antipsychotic drug in the management of treatment-resistant depression: a systematic review and economic evaluation. Health Technol Assess 2013;17(54)
43. Segal, Z., Pearson, J., & Thase, M. (2003). Challenges in preventing relapse in major depression. Journal Of Affective Disorders, 77(2), 97-108. doi: 10.1016/s0165-0327(02)00112-x
44. Borges S, Chen Y-C, Laughren TP, et al. Review of maintenance trials for major depressive disorder: a 25-year perspective from the US Food and Drug Administration. J Clin Psychiatry. 2014;75(3):205–214.
45. Durgam, S., Gommoll, C., Migliore, R., Chen, C., Chang, C., Aguirre, M., & Thase, M. (2018). Relapse prevention in adults with major depressive disorder treated with vilazodone. International Clinical Psychopharmacology, 33(6), 304-311. doi: 10.1097/yic.0000000000000236
46. Morrens et al, Increased use of augmentation therapies illustrates problematic management of treatment resistant depression. Preliminary results from a real-world evidence registry in Europe, the 54135419DEP4001 study. ISPOR November 2019.
47. Sapin, C., Fantino, B., Nowicki, M., & Kind, P. (2004). Usefulness of EQ-5D in Assessing Health Status in Primary Care Patients with Major Depressive Disorder. Health And Quality Of Life Outcomes, 2(1), 20. doi: 10.1186/1477-7525-2-20
48. Mago, R., Fagiolini, A., Weiller, E., & Weiss, C. (2018). Understanding the emotions of patients with inadequate response to antidepressant treatments: results of an international online survey in patients with major depressive disorder. BMC Psychiatry, 18(1). doi: 10.1186/s12888-018-1625-y
49. Mago, R., Fagiolini, A., Weiller, E., & Weiss, C. (2018). Healthcare professionals' perceptions on the emotional impact of having an inadequate response to antidepressant medications: survey and prospective patient audit. Annals Of General Psychiatry, 17(1). doi: 10.1186/s12991-018-0189-z
50. Mrazek, D., Hornberger, J., Altar, C., & Degtiar, I. (2014). A Review of the Clinical, Economic, and Societal Burden of Treatment-Resistant Depression: 1996–2013. Psychiatric Services, 65(8), 977-987. doi: 10.1176/appi.ps.201300059
51. Janssen. Data on File. 2020
52. Mauskopf, J., Simon, G., Kalsekar, A., Nimsch, C., Dunayevich, E., & Cameron, A. (2009). Nonresponse, partial response, and failure to achieve remission: humanistic and cost burden in major depressive disorder. Depression And Anxiety, 26(1), 83-97. doi: 10.1002/da.20505
53. Wright BM, Eiland EH, 3rd, Lorenz R. Augmentation with atypical antipsychotics for depression: a review of evidence-based support from the medical literature. Pharmacotherapy. 2013;33(3):344-59.
54. Posternak, M., & Zimmerman, M. (2007). Therapeutic effect of follow-up assessments on antidepressant and placebo response rates in antidepressant efficacy trials. British Journal Of Psychiatry, 190(4), 287-292. doi: 10.1192/bjp.bp.106.028555
55. Gastaldon, C., Raschi, E., Kane, J., Barbui, C., & Schoretsanitis, G. (2020). Post-Marketing Safety Concerns with Esketamine: A Disproportionality Analysis of Spontaneous Reports Submitted to the FDA Adverse Event Reporting System. Psychotherapy And Psychosomatics, 1-8. doi: 10.1159/000510703
56. Spravato® REMS (Risk Evaluation and Mitigation Strategy). Available at: https://www.spravatorems.com. Accessed 4 September 2020.
57. US Food and Drug Administration. FDA Adverse Event Reporting System (FAERS) public dashboard). Available at: https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard. Accessed 4 September 2020.
58. Toki T, Ono S. Spontaneous reporting on adverse events by consumers in the United States: an analysis of the Food and Drug Administration Adverse Event Reporting System database. Drugs Real World Outcomes 2018;5(2):117-128. doi: 10.1007/s40801-018-0134-0.
59. Feldman RL, Dunner DL, Muller JS, Stone DA. Medicare patient experience with vagus nerve stimulation for treatment-resistant depression. J Med Econ. 2013;16:62-74. [Seq.0028/Mod5.4/khan-a-2018-p-1455-1462]
60. IBM® MarketScan® Commercial Claims & Encounters (CCAE).

61. Kumar A, Bunker MT, Aaronson ST, Conway CR, Rothschild AJ, Mordenti G, et al. Durability of symptomatic responses obtained with adjunctive vagus nerve stimulation in treatment-resistant depression. Neuropsychiatr Dis Treat. 2019;15:457-68.

62. Vergunst FK, Fekadu A, Wooderson SC, Tunnard CS, Rane LJ, Markopoulou K, et al. Longitudinal course of symptom severity and fluctuation in patients with treatment-resistant unipolar and bipolar depression. Psychiatry Res. 2013;207(3):143-9.

63. Fekadu A, Wooderson SC, Rane LJ, Markopoulou K, Poon L, Cleare AJ. Long-term impact of residual symptoms in treatment-resistant depression. Can J Psychiatry. 2011;56(9):549-57.

64. Keller, M. (1992). Time to Recovery, Chronicity, and Levels of Psychopathology in Major Depression. Archives Of General Psychiatry, 49(10), 809. doi: 10.1001/archpsyc.1992.01820100053010.

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal Committee is interested in receiving comments on the following: • has all of the relevant evidence been taken into account? • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? • are the provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced.
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	SANE
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
Name of commentator person completing form:	[redacted]

Comment number	Comments Insert each comment in a new row.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	Do not paste other tables into this table, because your comments could get lost – type directly into this table.
1	We cannot comment on whether all of the relevant evidence has been taken into account, or on whether the summaries of clinical and cost effectiveness are reasonable interpretations of the evidence. As regards whether the provisional recommendations are sound and a suitable basis for guidance to the NHS, we welcome the committee's conclusions that treatment-resistant depression has a negative effect on people, their families and carers; that the effectiveness of current treatments for the condition is limited; and that there is an unmet need for new treatment options. We hope that, given these conclusions, further work will enable the committee to recommend esketamine for use in the NHS. In undertaking this work, we hope the committee will give due weight to the evidence given by patient experts. We reiterate the points we made in our submission of July 2019 and our comments in response to the appraisal consultation document issued in January 2020. Since we made those submissions, we have seen the effects on mental health of the Covid-19 pandemic. People with depression contacting us on our helpline have told us that they are being severely affected by the pandemic, experiencing a deterioration in their mental health and expecting a further toll with the continuation of restrictions and mental health services not always easy to access. For those with treatment-resistant depression, we fear there could be a much more deleterious effect, putting people at risk of self-harm and suicide. We believe this makes it all the more important and urgent that there be the most effective treatment response for those living with treatment resistant depression.
2	We know that for those living with treatment-resistant depression, there is a loss of hope that it can improve, or that any treatments might be helpful or effective. The expert patient evidence in the appraisal consultation document highlights this, and how the feelings of hopelessness increase when multiple courses of treatment do not work. We believe the fact that esketamine can have an effect within 24 to 48 hours of being administered, potentially saving patients the weeks or months of uncertainty that can be experienced with other anti-depressants, is critically important in offering hope to those for whom other treatments have not proved effective, and in seeking to alleviate the negative effect of the condition on patients and carers.
3	We are pleased that the committee concluded that the biological mechanism of esketamine could be innovative. Patients are having to rely on medications that are over 30 years old, which can have unpleasant side effects and do not work for everyone. Allowing esketamine as an additional treatment would offer the possibility of relief from suffering, and widen patient choice.
4	
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise and all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a 2nd version of your comment with that information replaced with the following text: 'academic / commercial in confidence information removed'. See the Guide to the processes of technology appraisal (section 3.1.23 to 3.1.29) for more information.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the appraisal consultation document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal Committee is interested in receiving comments on the following: • has all of the relevant evidence been taken into account? • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? • are the provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	British Association for Psychopharmacology
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Nothing to disclose
Name of commentator person completing form:	Form submitted by [redacted] Comments received from [redacted]

Comment number	Comments
	These comments have been answered to address the points below: • has all of the relevant evidence been taken into account? yes • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? no • are the provisional recommendations sound and a suitable basis for guidance to the NHS? no

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	We are concerned that this recommendation may imply that ……..
1	From NICE: "But it is unclear how effective esketamine is because of the way the trials were done." In discussing the trials, the committee noted that TRANSFORM 1 and TRANSFORM 3 did not show significant results. BAP response: We note that nasal esketamine has been deemed an effective medication by 2 major regulatory authorities (FDA and EMEA) which have approved it for patients with Treatment Resistant Depression. It is important to note that Transform -2 allowed flexible dosing of esketamine, which is in line with the relevant Summary of Prescribing Information. However, the regulatory approval was based on data from 2 positive phase 3 studies (studies 3002, Transform 2 and 3003, Sustain 1), as well as supportive data from additional phase 2 and 3 studies, demonstrating consistent efficacy of esketamine and evidence supporting long-term safety (study 3004). To approve a new antidepressant, our understanding is that regulatory authorities generally require two positive, short-term, adequate and well-controlled studies to meet the (regulatory) standard for substantial evidence of effectiveness. Randomized withdrawal studies are typically conducted after approval to support an additional maintenance claim. For esketamine, however, regulatory authorities required both short- and long-term data in the initial application due to the novelty of the product. It should also be noted that, when the results of TRANSFORM 1 and 2 are pooled, the results do indicate a significant effect. Therefore, we would conclude that for adults aged 18-65 years significance was demonstrated.
2	From NICE: "The response and remission evidence from TRANSFORM-2 should be considered with caution because of the short duration of the trial" The committee took into account a Consultee who stated this time period had little bearing on the treatment for depression, though no evidence is given for this statement. BAP response: We do not understand this point raised by the committee. It is unclear why the committee should choose to reference the NICE guideline on depression and antidepressant treatment, given that esketamine is a different class of drug, with a different mode of action-and effects on depression scores, with separation from placebo early on in treatment. If anything, effects at 4 weeks are likely to be an underestimate of overall effect, due to the fact that response and remission rates would not likely decrease in value both in the interventional and the control arm (but would only likely increase) with a longer observation period. In addition, study 3003 demonstrated a statistically significant long-term effect of esketamine plus oral antidepressant in maintaining a state

	of remission or response when compared against oral antidepressant alone. An additional point to make about the duration of trials is that 28 days may seem short for trials of antidepressants that are typically taken for many months, but this is the internationally agreed frame for a licensing trials because of the ethical difficulty of leaving people on placebo for longer periods. Therefore, it is inevitable that prelicensing longer term studies are open label, and that post licensing studies are used to clarify longer effects. We appreciate the NICE committee's frustration with this, but it is the reality of research into all new drug treatments for depression and it does not seem reasonable to withhold a drug from the widespread use that is necessary to obtain the long term information. The overall evidence clearly does not support the logic implied in the comment, that the results of a 4-week duration trial has no impact on real-world clinical significance.
3	From NICE: The TRANSFORM-2 study is not powered to detect difference in effect between treatment arms so could show a false positive result BAP response: The committee rightfully point out that in this trial a higher than normal placebo response than would be expected was seen. They also highlight potential regression to the mean. Both of these aspects of trial design mean that any placebo/drug difference would be minimised. Therefore, (in our view) the correct interpretation is that these data probably underestimate the treatment effect. Furthermore, the committee also point to the short, 4-week duration-another factor that would minimise the drug/placebo difference. The fact that the initial power calculation was based on a higher estimated difference does not seem relevant here, as the effect size is reported in the trial.
4	From Nice: "Withdrawal effects are difficult to distinguish from symptoms of depression" The report quotes a consultee who queries whether withdrawal from esketamine could confound relapse rates. No evidence is given. BAP response: The evidence that exists with regards to withdrawal effects is from people who misuse ketamine, and of the two clinical reports in the literature (both from reviews), less than 50% of ketamine abusers developed withdrawal symptoms. These people were using ketamine daily, at doses up to 9g-far in excess to that used in the esketamine trials. Furthermore, there is very little information on cardinal features of withdrawal in these reports. In Study TRD3003, although there was a high number of relapses in the first month in those switched to placebo nasal spray, it is unlikely that a pharmacologic withdrawal effect contributed given that the decrease in esketamine plasma concentrations is rapid for the initial 2 to 4 hours and more gradual thereafter (with a mean terminal half-life, 7-12 hours), with steady state never reached with intermittent dosing. Moreover, this high rate of early relapse is similar to that observed after cessation of electroconvulsive therapy. There are no known rebound effects after electroconvulsive therapy discontinuation. The high rates of early relapse after esketamine discontinuation and those observed by Rush et al for patients in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study at level 3 or 4 (i.e., who had failed 2 and 3 prior antidepressant treatments, respectively) more likely reflect a greater vulnerability to relapse among patients with TRD during maintenance treatment with an antidepressant alone. The FDA report states that "Acute esketamine withdrawal is likely not a factor, as dosing is infrequent during the maintenance phase." It is physiologically implausible for such infrequent dosing to cause a withdrawal syndrome. Therefore, we conclude that there is nothing evident to us to suggest a withdrawal syndrome. Furthermore, no evidence is presented to show that items on the withdrawal checklist correlate with

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	those of the MADRS, and the trial authors themselves state, "No evidence of a distinct withdrawal syndrome was observed during the 2 weeks after cessation of esketamine nasal spray as assessed by the 20 item Physician Withdrawal checklist".
5	From NICE: The differences in relapse rate in the SUSTAIN-1 trial data should be considered with caution BAP response: This argument was initially put forward in a comment on Lancet Psychiatry (Lancet Psychiatry. 2019; 6: 977-979) and we note has since been addressed by the company (Lancet Psychiatry VOLUME 7, ISSUE 3, P232-235, MARCH 01, 2020): neither the company nor the FDA (after site inspection) found any reason to exclude data from the site in Poland which is the subject of the author's comment. Nonetheless, a sensitivity analysis was performed excluding this site and using a statistical method appropriate for time to event data. Statistical significance was maintained (log-rank test $p<0.05$) and the results remain consistent with the primary efficacy analysis. It is puzzling why the committee should continue to discuss this point.
6	From NICE: Healthcare resource use costs should be made equal across both arms in the current model. BAP response: • We feel that it is clinically more appropriate to assume that healthcare resource use and hence medical costs is/are higher when patients with TRD are symptomatic compared to when patients are in remission than to assume (as NICE appears to) that healthcare resource use and medical costs are the same, independent of clinical effectiveness of the treatment received and the health state of patients.
7	From NICE: The effect of subsequent treatments is underestimated, and the ERG's adjustment is more plausible BAP response o It is important to highlight that the majority of patients who have not responded to many antidepressant treatments have very poor outcomes in the long-term o This is about the total cohort of patients with TRD, and not a subgroup of patients who have received very intense treatment in an extremely specialised hospital setting (which has since closed) and were doing well after their intense treatment as shown in the Fekadu and Wooderson study

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise and all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a 2nd version of your comment with

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

that information replaced with the following text: 'academic / commercial in confidence information removed'. See the Guide to the processes of technology appraisal (section 3.1.23 to 3.1.29) for more information.

- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the appraisal consultation document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal Committee is interested in receiving comments on the following: • has all of the relevant evidence been taken into account? • are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? • are the provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	The Royal College of Psychiatrists
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

Name of commentator person completing form:	[redacted], [redacted] and [redacted] **Commercial:** [redacted] **Academic**: [redacted]

Comment number	Comments Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
1	**Has all of the relevant evidence been taken into account?** Evidence on the efficacy and safety of ketamine should be considered. The neuromodulatory effects of esketamine in the brain are likely to be identical or extremely similar to ketamine at equivalent doses. In this regard, the highest level of impartial evidence is likely to be Cochrane Reviews. These show that evidence in support of ketamine in depression is generally of poor quality, involving small samples, and with efficacy only shown over brief (clinically likely to be irrelevant) time periods. Also, risk of bias was often unclear, due to a lack of reporting. See: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011612.pub2/full

Please return to: **NICE DOCS**

	?highlightAbstract=ketamine%7Cketamin https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011611.pub2/full?highlightAbstract=ketamine%7Cketamin Some academic institutions have patents and other intellectual property regarding ketamine and/or esketamine. This is more a case in the USA. Such conflicts of interest are often not mentioned when members of those institutions give presentations as to the apparent benefits of ketamine and/or esketamine.
2	**Comment on Esketamine is likely to be used later in the treatment pathway because it has a higher treatment burden than other treatments** We agree that the position of esketamine in the treatment pathway is initially likely to be at least fourth or fifth line – i.e. after trials of augmentation. We also agree that for some patients this is because of the burden of treatment. Patients may not drive following esketamine treatment until they have had restful sleep. They can return home using public transport when they are fully recovered. However, we consider that this later use also reflects its expense, novelty and association with a drug of abuse, more than the clinical data. Compared with the alternatives it is not obviously less safe. Therefore, particularly once costs come down, and particularly for patients are well supported, it is likely to be used earlier in the pathway. For some of those who are less well supported it may be more appropriate to provide better hospital or volunteer transport than to withhold the medication until later.
3	**Comment on the response and remission evidence from TRANSFORM-2 should be considered with caution because of the short duration of the trial** We agree that 28 days does seem short as a primary end point for trials of antidepressants that are typically taken for many months. However, we do not think it is right to say that this has 'little bearing' on the treatment of depression. This is the internationally agreed time frame for licensing trials because of the ethical difficulty of leaving people on placebo for longer. Uniquely amongst programmes for a new antidepressant, the short term 28-day data in TRANSFORM are supplemented by the high quality data of the 1 year study SUSTAIN 2. Usually, it is lower quality post licensing studies that are used to clarify longer effects. It would not seem reasonable to withhold this drug from widespread use on the basis of a criticism that can be levelled at all other antidepressants which are in current use. 'The committee acknowledged that splitting the data into 2 groups could have inflated the differences between arms, particularly because the mean reduction in MADRS was near to the threshold for response in both arms at day 28. So, people could meet the criterion for symptom response in 1 arm but only have minimal differences in MADRS score in the other arm'. We do not agree that splitting the data into 2 groups could have 'inflated the difference'. The fact that the difference in remission, which is based on an absolute threshold level of the MADRS, between the two arms in TRANSFORM 2 (21.5%) is greater than the difference in response (17.3%), which is dependent on change relative to baseline level, effectively disproves the possibility of an inflated effect. We would further make the point that response and remission are entirely conventional,

	pre-specified, measures. This new concept of a 'threshold in response' does not make sense when the difference in MADRS needed to meet criteria for response will vary for each participant depending on baseline. It is no more right to make decisions based on a NICE-generated post-hoc analysis which suggest that an effect size is 'near to threshold' than it would be to make decisions based on company-generated post hoc analyses which showed big effects.
4	**Comment on the TRANSFORM-2 study is not powered to detect difference in effect between treatment arms so could show a false positive result** We do not understand why universally accepted standards for accepting a difference between two arms of a trial, are described as potentially a 'false positive'. It is of course possible that any result could be a 'false positive', but this is why we have accepted norms of statistical significance. The language here seems inappropriate. One would not accept a comment from a company which asserted that their non-significant result was potentially a false negative because a study was underpowered! The powering of the study is based on the number of patients to detect a difference assuming a specific degree of variance. It is possible that the difference was statistically significant despite the smaller than estimated effect size because, even though the difference was smaller, the degree of variance was lower.
5	**Comment on withdrawal effects are difficult to distinguish from symptoms of depression** We agree that it is difficult to conclusively disprove that a new symptom arose because of stopping the drug rather than because relapse. However, the pattern of new symptoms provide important evidence as to which was happening and this does not appear to have been considered. In SUSTAIN 2 (Wajs et al 2020 Supplementary 5) the following effects were common (all >20% in the second week after cessation): insomnia, anxiety-nervousness, dysphoric mood-depression, fatigue – lethargy – lack of concentration, irritability, difficulty in concentration. These are all symptoms of major depressive disorder. By contrast the following symptoms were much less common: loss of appetite, nausea -vomiting, diarrhoea, poor coordination, sweating, tremulousness, dizziness-lightheadedness, headache, muscle stiffness, weakness, increased acuity sound smell touch, paraesthesias, depersonalisation-derealisation. With the exception of loss of appetite, these are not features of major depressive disorder. The dominant problem is therefore more likely to be relapse in depression rather than new symptoms occurring due to a change in physiology induced by the drug. Increased feeling of hopelessness on withdrawal are an important problem, but are much more likely to be due to relapse in depression rather than being caused by the drug. The short acting nature of the drug means that if it did induce some sort of change of physiology which caused withdrawal symptoms, then these effects would be expected to occur between each weekly dose (thereby undermining its beneficial effect), rather than after the end of a course. This was not observed. For these reasons, we think the results of SUSTAIN 1 should be taken at face

	value. The main implication of SUSTAIN 1 is that the drug needs to be taken continuously to prevent relapse. It undermines the company's assertion, made on the basis of much less direct evidence, that relapse will not occur if it is withdrawn later.
6	**Comment on the differences in relapse rate in the SUSTAIN-1 trial data should be considered with caution** There seems to be a disparity between the conclusion – that the results of SUSTAIN 1 should be treated with caution – and the text which follows, all of which seems to point to reasons why the data of an outlier should not be excluded. The choice of language here seems inappropriate.
7	**Comment on the evidence for esketamine is limited in its generalisability to the NHS** **Severity** We agree that the trial data are limited in the degree of generalisability to populations that are more severe, but do not think this is a strong argument against adoption. Current practice is to use the same antidepressants in people with depression of all severities. The choice of antidepressants at different points on the treatment pathway is determined by side effect profile rather than by different antidepressants having different efficacy in different severities. Comorbidities The poor generalisability associated with the exclusion of patients with comorbidities is also relevant, both to safety and efficacy. Most psychiatric disorders, are associated with depression; and each will sustain and fuel the other. This is a contributory factor to high rates of prescribing of antidepressants in the population. We think the appropriate way to manage the risks of prescribing in patients with comorbid illness is through good phase 4 studies following adoption, rather than by withholding the drug from people with 'pure' resistant depression because it might be used in people with complicating comorbidities. This can work in unexpected ways. For example, there are data from multiple studies suggesting that ketamine can be of benefit in reducing substance misuse. Clearly, however, there are also risks in people who are vulnerable to developing addiction, as reflected in the datasheet.
8	**Comment on it is not appropriate to adjust the efficacy estimates of the placebo arm in the trials** Whilst we agree with the company's assessment of the influences on the placebo effect, we agree with the committee that the sort of post-hoc adjustment which the company applied was not appropriate.
9	**Comment on safety must be considered when administering and monitoring esketamine** We agree with the committee that a registry is required. Further, we consider that such a registry should be interrogatable. Otherwise, those wishing to prescribe other rapidly acting antidepressants (eg IV ketamine) cannot be sure whether an individual is additionally taking esketamine nasal spray and is 'topping up'.

	This would be a first step on the way to the use of systems such as Safescript (now mandatory in Australia) and Drug Prescribing Monitoring Programmes (as used in every state of the US) which require that, before prescribing, doctors intending to prescribe certain scheduled drugs must interrogate databases to ascertain existing and previous scheduled drug use. We agree with the committee and the regulators that the signal is not, at present, strong enough to justify withholding the drug from the larger number who may benefit. An interrogatable registry will help in tracking the extent of suicidal behaviour associated with relapse or non-response to esketamine. This is a phase 4 task.
10	**Comment on economic model** 3.17 – 3.19 The company's economic model does not reflect the course of the disease We agree that there are 'minimal long-term outcome data for people with treatment-resistant depression' to inform modelling. The higher cost of the drug makes this more important than it is for other cheaper oral antidepressants.
11	**Comment on the effect of subsequent treatments is underestimated and the ERG's adjustment is more plausible** We agree with the committee that clinical practice would not be that 3 treatments would be attempted within 12 weeks as each successive treatment failed. Cycling between treatment takes much longer than this.
12	**Comment on the cost of a course of esketamine treatment may be underestimated** The committee is concerned about variations in the dose and frequency of treatment. This data already exists. The company's data, as submitted to the FDA, shows that a higher proportion of those who remit but do not respond take maintenance esketamine at the shorter, weekly maintenance interval (69%) than those who remit (34%). In other words, those who respond less well take it more frequently.
13	**Comment on A 1 to 2 ratio of nurses to patients is an appropriate resource cost during post-administration monitoring** We disagree slightly with the committee here. Based on the experience of the 5 UK centres which administer IV ketamine - for which the recovery time and requirements are likely to be similar if not slightly higher than for nasal esketamine - we consider that a 1 to 3 ratio more accurately reflects the need for healthcare staff supervision. Post treatment observation can be done by a healthcare assistant and, depending on the layout of the clinic needs only to be intermittent rather than continuous. It does not require a qualified nurse. We agree with the clinical expert that the staffing need will change as clinics develop experience and efficiency of procedures. A typical ECT department would be able to start by treating esketamine patients at the end of their twice weekly ECT lists, thus avoiding employment of new staff until numbers justified a new bespoke clinical session. Fairly quickly, a single nurse and healthcare assistant can run a clinic with 3 concurrent patients each of which will be in clinic for about 2 hours in total. In a clinic which has the beds/chairs to manage 3 simultaneous patients, two staff would

	be comfortably able to treat 6 patients in a session, including time for recording notes and, depending on its complexity, completing the registry. It is important to note that, as with directed observed administration of other CDs, a doctor does not need to be immediately present for the treatment. [redacted]
14	**Comment on significant investment will be needed to use esketamine in the NHS, but costs are difficult to quantify** Based on our experience, we think the only physical infrastructure likely to be required in an ECT suite is a Controlled Drug cabinet. In other settings it may also be necessary to purchase suitable comfortable chairs. The processes for transporting drugs to the ECT exist already and are part of routine hospital transport systems so this does not incur new costs. The arrangements for disposal of used devices consist of putting a bespoke bin (like a large blue sharps bin) in the department which, when full, is transferred back to pharmacy for formal disposal of the remnants of the devices. This uses existing transport arrangements and again is low cost. Training: The procedure is not complex, training materials are provided by the company and this could be accomplished within existing allocation of training time.
15	**Comment on it will take time and resource use for esketamine to become part of clinical practice** We agree that esketamine is potentially disruptive to existing practice but observe that this may be a good thing. For example, patients with resistant depression commonly find that they become disillusioned with CMHT services because, however good the support, their condition does not change (by definition). When they have a treatment which abruptly helps, their care rapidly aligns with the service which provides it. In our experience, this commonly then results in the CMHT wishing to discharge the patient. The service providing esketamine then finds itself with a rapidly increasing caseload of patients who, if they relapse, are potentially at high risk. One way of managing this risk is to have shared care with the CMHT, but this duplicates effort and can seem pointless to the patient. A better solution may be to draw the resource into the new service from the old. This sort of disruption is to be welcomed – but, like all disruption, may initially be unpopular. We agree that esketamine services should not be confined to ECT services and that community settings would be suitable. However, the infrastructure – a clinic with comfortable chairs, separated by curtains, which is suitable for administration and recovery - is common in many NHS settings. We also agree that the reality of NHS processes is such that the lead times of 6-12 months for implementation quoted are realistic. However, this is driven by institutional barriers to introducing new technologies. Because the 'technology' is very simple, private clinics will be much quicker in set up. In conclusion, we would not describe the costs of setting up a clinic as 'substantial'. The

Esketamine for treatment-resistant depression [ID1414]

Consultation on the appraisal consultation document – deadline for comments 5pm on 25 September 2020 **email:** NICE DOCS

	staff running costs could reasonably be estimated as a third of a session of a band 6 and a band 3 nurse – about £20 per patient per treatment.

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise and all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a 2nd version of your comment with that information replaced with the following text: 'academic / commercial in confidence information removed'. See the Guide to the processes of technology appraisal (section 3.1.23 to 3.1.29) for more information.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the appraisal consultation document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Comments on the ACD received from the public through the NICE Website

Name	[redacted]
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
Patients suffering from treatment resistant major depression use more healthcare resources in terms of hospital admissions, GP consultations and psychological treatments than patients who have responded to treatment and recovered. It is vitally important to have another helpful and useful treatment other than ECT that patient can access. E

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
There are Clinical Treatment Teams as well as ECT Suites in some areas, which can form the spine for setting up the clinics. They have resuscitation and monitoring equipment as well as appropriately trained professionals. This consideration will give a better interpretation of cost-effectiveness.

Are the recommendations sound and a suitable basis for guidance to the NHS?
The conclusions made by NICE at this stage of lack of cost-effectiveness is misjudged. It is in the interest of the NHS and our patients to grant approval for use in combination with an SSRI or SNRI as a third line choice. The goal must be to improve the patient's quality of life and where possible a return to gainful employment.

Name	[redacted]
Comments on the ACD:	

General Comment:
As someone with treatment resistant depression I need more options. I have frequent suicidal thoughts and attempts. I have heard first-hand accounts of people in the US who have benefited from esketamine and I want that opportunity to be relieved from my depression. Depression affects so many people in my life, not just me. My husband, family and friends. SSRIs are limited and almost impossible to come off. We need other options. The money that this will cost may save thousands of lives. If this was a new cancer treatment we would be endorsed.

Name	[redacted]
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
Not sure

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
No - see comments

Are the recommendations sound and a suitable basis for guidance to the NHS?
No - see comments

General Comments:
We support the careful approach that you have taken to considering approval of this new treatment for depression, especially given its use as an anaesthetic agent and party drug with recognised abuse potential and association with considerable harm.

We agree with the position of the committee that it does not seem to be simply a question of whether esketamine is cost-effective but whether it is effective at all. Depression is a long-term condition that can make people's lives harder. It therefore does not seem particularly relevant what the effects of any drug are after 4 weeks. This may be an appropriate time point for assessing response to an infection but not to mood states. There are a number of substances that might temporarily improve mood – like opioids, benzodiazepines, alcohol, cocaine and many other recreational substances- due to their effect of inducing a euphoric state. However, it is quite a different question as to whether these substances will produce a change that will be beneficial to a person in the long-term.

We fear that short-term studies may demonstrate effects that are not borne out in the long-term – similar to that for many illicit substances, for which the long-term outcomes are generally dysphoric states. Indeed, this is the case for ketamine users who are general found to be dysphoric, even after they stop the drug (Morgan & Curran, 2012). It is therefore imperative that drugs which are provided in the NHS are rigorously tested for their long-term effects – both positive and negative.

In addition to this, the effects of esketamine do not seem established even in the very short-term horizon of 4 weeks. Only one out of three short-term trials showed a significantly statistical effect on depression scores, and the effect was very small, with many people pointing out that it does not register as a clinically significant effect (C. Gastaldon, Papola, Ostuzzi, & Barbui, 2019; Horowitz & Moncrieff, 2020). We also understand that in two further as yet unpublished trials of esketamine in suicidal patients that the effect on depression scores was also not significant. It seems very unconvincing that this drug has any positive effects even in the short term. It seems rather concerning that although the company performed a 24 weeks study of placebo versus esketamine that they did not report the outcomes of this study in terms of depression scores. One wonders whether this is due to the results not aligning with their commercial objectives. Moreover, the effects of human contact seem to grossly outweigh the effects of esketamine in the trials. A 17-point reduction on the MADRS scale derived from salt water

spray and time with a nurse seems to us a considerable effect, dwarfing the effect due to the drug (up to 4 points in some studies).

We understand that it is highly unusual for a discontinuation study of a drug to be included in proving a drug works when the drug is known to cause dependence and withdrawal as ketamine is. This study design would be deliberately confusing because withdrawal effects from esketamine would mimic symptoms of 'relapse' and so make coming off the drug look detrimental. We also understand that the company did not report withdrawal effects in its results meaning that it is not clear that this possibility was excluded. It would seem that the results from this study are not reliable as a result. Even not taking this into account, the difference between continuing or discontinuing do not seem to be very different at all after a few weeks suggesting that the drug is not really effective at preventing 'relapses.'

Lastly the harms of this drug have been downplayed, but there is good reason to think that they could be quite significant. Ketamine is known to cause a number of health issues in recreational users or in patients given it in anaesthetic doses (admittedly larger doses than employed her, but notably patients will be given esketamine much more often than in anaesthetic practice). It is known to cause ketamine bladder, whereby the bladder wall is worn away over time, leading to people needing catheters to pass urine. It can cause heart attacks, and strokes due to the increase in blood pressure ('spikes'). It can cause motor vehicle accidents because ketamine has profound effects on hand-eye co-ordination, judgement and decision making. It has also been associated with suicides. This may be due to the psychotic symptoms(Beck et al., 2020; Wood et al., 2011) it is known to cause at sub-anaesthetic doses (including the doses tested by Janssen). It may be due to withdrawal effects from the drug (Schatzberg, 2019). It is hard to know the exact reason but it is surely very concerning that all these events occurred in the esketamine arm of Janssen's studies more frequently (and sometimes exclusively, in the example of suicides) in the esketamine arm compared with the placebo arm (C. Gastaldon et al., 2019; Horowitz & Moncrieff, 2020). It is perhaps more concerning to see these trends for harms extend into real-world practice with the same group of harms occurring in US in the year since the drug has been approved for use (Chiara Gastaldon & Kane, 2020). It is also alarming that the doses of esketamine used in clinical trials have also been shown to alter neurodevelopmental pathways in animal models leading to severe cognitive and behavioural impairments(Zimmermann, Richardson, & Baker, 2020). The long-term effect on the adult brain has not been investigated but these results are foreboding.

This combination of factors – a lack of clear efficacy in the short-term, a lack of evidence of benefits in the long-term, very serious signals about harms from this drug, known risks of abuse and misuse – makes it seem reasonable to err on the side of caution and await more robust proof of efficacy, especially in the long-term (of a year or more) and verify that the safety signals are not likely to increase morbidity and mortality of users of esketamine (or other road users).

It is also striking that the academics who have come out of support of this drug as having a ‘novel action’ or a ‘breakthrough’ are all paid by Janssen. For example, in a recent letter to the British Journal of Psychiatry (Kasper, Young, Vieta, Goodwin, & Meyer-Lindenberg, 2020), echoing the arguments put by Janssen to other critical papers, and echoing some of the submissions made to the NICE consultation process, all five authors receive money from Janssen. At least two authors are principal investigators on esketamine studies funded by Janssen. Their remarks about the ‘novelty’ of esketamine are reflected in submissions to NICE as outlined in the slides from the recent consultation hearing. The British Association of Psychopharmacology (BAP) has also made comments very supportive of the drug, without being supported by evidence. It has not been made clear whether the BAP has received direct payments from Janssen. Lastly, the clinical expert makes a number of points in the consultation document that are repeatedly favourable to Janssen’s position (eg that a 4 point improvement on the MADRS represents a clinically significant difference, in contradiction of the existing evidence). There are likely to be few people with experience of ketamine in the UK given its limited deployment in clinical practice and many of the experts, such as Rupert McShane or Hamish McAllister-Williams (https://mood-disorders.co.uk/admin/resources/hamish-mcallister-williamsvns-and-restore-life17sept18.pdf), have close relationships with the manufacturer, including direct payments for consultancy as well as research support. We wonder whether the financial connections to the manufacturer of clinical experts involved in the committee’s deliberations or that have been called on to give expert testimony might have influence the opinions presented, especially when unsupported by the existing research evidence but consistent with the manufacturer’s commercial objectives. Furthermore, it has been recognised that drug manufacturers often use the small Scottish market to put pressure on NICE to generate a favourable review in order to get access to the more lucrative English market. This seems to be occurring in this case with Janssen where they have agreed to subsidise esketamine in Scotland to render it ‘cost-effective.’ We hope that such political machinations will not influence the committee’s appraisal of the scientific evidence.

On the other hand analysis of esketamine by experts independent of financial ties to the manufacturer has unanimously concluded that the drug is not effective and its safety has been questioned (Cristea & Naudet, 2019; C. Gastaldon et al., 2019; Chiara Gastaldon & Kane, 2020; Horowitz & Moncrieff, 2020; Schatzberg, 2019; Turner, 2019). It is also notable that the national health evaluators in France, Denmark and Sweden have generally given a negative evaluation of esketamine’s usefulness and not approved it for widespread use (although some countries have approved it as a fourth line antidepressants in some cases).

Overall, given a lack of evidence for the effectiveness of this medication, a lack of long-term data and worrying danger signals, it would seem mandatory to demand greater evidence of safety and effectiveness before approving this potentially harmful treatment for widespread use.

Sincerely,

[redacted], Clinical Psychologist, [redacted]

[redacted], Psychiatrist, [redacted]

[redacted], Unaffiliated Psychiatrist, Oxford, England

[redacted], Psychiatrist (retired)

[redacted], Consultant Psychiatrist, [redacted]

References

Beck, K., Hindley, G., Borgan, F., Ginestet, C., McCutcheon, R., Brugger, S., … Howes, O. D. (2020). Association of Ketamine With Psychiatric Symptoms and Implications for Its Therapeutic Use and for Understanding Schizophrenia: A Systematic Review and Meta-analysis. JAMA Network Open, 3(5), e204693–e204693. https://doi.org/10.1001/jamanetworkopen.2020.4693

Cristea, I., & Naudet, F. (2019). US Food and Drug Administration approval of esketamine and brexanolone. The Lancet Psychiatry, 1–3. https://doi.org/10.1016/ S2215-0366(19)30292-5

Gastaldon, C., Papola, D., Ostuzzi, G., & Barbui, C. (2019). Esketamine for treatment resistant depression: A trick of smoke and mirrors. Epidemiology and Psychiatric Sciences, 08(Ci), 91–94. https://doi.org/10.1017/S2045796019000751

Gastaldon, Chiara, & Kane, J. M. (2020). Post-Marketing Safety Concerns with Esketamine : A Disproportionality Analysis of Spontaneous Reports Submitted to the FDA Adverse Event Reporting System. https://doi.org/10.1159/000510703

Horowitz, M. A., & Moncrieff, J. (2020). Are we repeating mistakes of the past? A review of the evidence for esketamine. The British Journal of Psychiatry, 1–4. https://doi.org/10.1192/bjp.2020.89

Kasper, S., Young, A., Vieta, E., Goodwin, G., & Meyer-Lindenberg, A. (2020). Letter to BJPsych in response to Horowitz and Moncrieff. British Journal of Psychiatry. https://doi.org/https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/are-we-repeating-mistakes-of-the-past-a-review-of-the-evidence-for-esketamine/C4BDC70050164FD9D88DF40967C2853A#fndtn-comments

Morgan, C. J. A., & Curran, H. V. (2012). Ketamine use: A review. Addiction, 107(1), 27–38. https://doi.org/10.1111/j.1360-0443.2011.03576.x

Schatzberg, A. F. (2019). A word to the wise about intranasal esketamine. American Journal of Psychiatry, 176(6), 422–424. https://doi.org/10.1176/appi.ajp.2019.19040423

Turner, E. H. (2019). Comment Esketamine for treatment-resistant depression : seven concerns about efficacy and FDA approval. The Lancet Psychiatry, 1–2. https://doi.org/10.1016/S2215-0366(19)30394-3
Wood, D., Cottrell, A., Baker, S. C., Southgate, J., Harris, M., Fulford, S., … Gillatt, D. (2011). Recreational ketamine: From pleasure to pain. BJU International, 107(12), 1881–1884. https://doi.org/10.1111/j.1464-410X.2010.10031.x
Zimmermann, K. S., Richardson, R., & Baker, K. D. (2020). Esketamine as a treatment for paediatric depression: questions of safety and efficacy. The Lancet Psychiatry, 0366(19), 10–11. https://doi.org/10.1016/s2215-0366(19)30521-8

Name	[redacted]
Comments on the ACD:	

General Comment:
As an ex Expert Clinical Assessor for MHRA, I fully support not approving esketamine. The reasons are laid out in a BMJ essay: The trouble with antidepressants: why the evidence overplays benefits and underplays risks--an essay by John B Warren
http://bmj.com/cgi/content/full/bmj.m3200

Name	[redacted]
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
In clinical practice, it takes a substantial amount of time (measured in years rather than months) for a patient to trial (at a therapeutic dose) the currently available different types of oral anti-depressants before ECT might be considered. This leaves a gap in the available depression treatment care pathway. New treatments for treatment-resistance depression are therefore urgently needed to reduce the burden of depression for the patient and carers.

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
In my experience as Neuromodulation Lead Nurse for [redacted] and [redacted] (National Association of Lead Nurses for ECT), the higher treatment burden claimed for Esketamine is no different (or possibly less) then the costs associated with ECT or Transcranial Magnetic Stimulation (TMS). Also, there are no significant cost issues (other than initial staff training) in using existing ECT clinics for Esketamine treatment. ECT clinics already have beds, monitoring equipment and staff in place. In addition, many other existing community mental health clinics (e.g. those set up for monitoring olanzapine depots) could be easily adapted to additionally monitor Esketamine treatment.

Are the recommendations sound and a suitable basis for guidance to the NHS?
The current conclusion does not currently take into account the inordinate length of time a patient must currently endure to receive appropriate treatment. If Esketamine can be proven to substantially reduce this time period then it should be considered for use in the NHS, especially as I believe that the costs estimates for Esketamine have been exaggerated by not including the considerable costs of ECT which is where such patients are often finally treated.

Name	[redacted]

Comments on the ACD:

Are the recommendations sound and a suitable basis for guidance to the NHS?
Could esketamine be referred to as esketamine nasal spray in the title of the technology appraisal, the recommendations and throughout the TA documents, so it is clear which esketamine product this TA is referring to. There are also esketamine injectable preparations available in the UK (which are licensed for various indications associated with anaesthesia).

Name	[redacted]

Comments on the ACD:

Has all of the relevant evidence been taken into account?
See below

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
See below

Are the recommendations sound and a suitable basis for guidance to the NHS?
See below

General Comments:
Thank you for your thoughtful consideration of the issues relating to the approval of esketamine. We offer some further points of clarification regarding the Appraisal consultation document:

3.1 Treatment-resistant depression.

The characterisation of ‘treatment-resistant depression’ as severe or burdensome is not necessarily accurate. ‘Treatment-resistance’ refers more to the drug treatment of depression than the nature of the condition itself (and is borrowed inappropriately from the concept of ‘antibiotic resistance’). Failing two antidepressants, which themselves have marginal

efficacy,(Jakobsen et al., 2017) (as explained in the BMJ recently)(Warren, 2020) is not necessarily evidence of a severe condition. Furthermore, the entry criteria for the Janssen studies excluded people with suicidal thoughts, past history of ECT and co-morbidities so it is not clear that esketamine has been tested in a group that corresponds to most clinicians' idea of 'treatment-resistant depression' or 'severe depression' entails. Most widely used definitions of 'treatment-resistant depression' include patients who have failed numerous different antidepressants from different classes (often explicitly including MAOIs, TCAs or SNRIs), have psychotic features, have trialled ECT (many staging models of 'treatment-resistant depression' take into account the number of sessions of ECT received), have trialled antipsychotics, and psychotropic augmentation strategies (Ruhé, van Rooijen, Spijker, Peeters, & Schene, 2012). Therefore, the patients included in the trial would not fit most clinicians' impression of what severe or 'treatment-resistant depression' constitutes. This impression is further emphasised by the fact that this group of patients showed a very significant improvement to placebo (17-point MADRS reduction). Even if the TRD concept does represent something more severe, experience with numerous other drugs shows that the treatment is likely to be extended to many people with less severe conditions.

3.2 Unmet needs.

Therapeutic hopelessness arises when pursuing ineffective treatments that have received hype from the marketing arms of their manufacturers, often echoed in views expressed by often-conflicted academics and similarly conflicted patient support groups (Fabbri, Lai, Grundy, & Bero, 2018; Fabbri et al., 2020; Lock, Seele, & Heath, 2016). Adding another ineffective medicine, associated with much manufactured hype will only contribute to hopelessness (see patient accounts below in Appendix). It is also a concern that patients have internalised 'treatment-resistance' as a self-descriptor, when it reflects treatment efficacy, rather than the patients themselves.

3.6 The effect of psychological therapy.

The clinical expert has suggested that "the lack of inclusion of psychotherapy in the studies of esketamine appears to be setting a hurdle that not a single currently recommended pharmacological treatment has surpassed." This may be true regarding regulatory approval, but with regards practice guidelines such as those set out by NICE, this is untrue – other classes of antidepressants have been carefully compared to psychological therapy. As psychotherapy is an intervention that is as effective as most drug treatments in the short term, but which some evidence suggests is more effective in the long-term and maintains its effects after the end of treatment, and has fewer side effects than drug treatments it would be more prudent to raise the bar for the evaluation of drug treatments to require psychotherapy as a comparison, rather than lower the bar for esketamine.

Although the clinical expert has suggested that some people have a condition which is too severe to be candidates for therapy, the fact that the group recruited for these studies showed a 17-point reduction in MADRS scores from passive social contact suggests that this would be a group that might respond to psychotherapy and that this comparison should be given consideration.

3.7 Efficacy of esketamine.

The committee recognised that there were two negative trials regarding the efficacy of esketamine compared with placebo. We now know there are a further three trials that demonstrate no statistically significant difference between placebo and esketamine: ASPIRE-1(Fu et al., 2019) and ASPIRE-2(Ionescu et al., 2019), and Canuso et al (2018)(Canuso et al., 2018). The ASPIRE trials were conducted by Janssen in suicidal patients meeting a diagnosis of MDD, and have so far only been presented as posters as conferences. They both found no statistically significant difference between placebo and esketamine groups at day 25 and no difference in suicidality at 24 hours, questioning the claim that esketamine is effective for suicidal thoughts (Fu et al., 2019; Ionescu et al., 2019)). Overall, there are therefore five negative trials (on MADRS score at day 25), compared to one short term trial with statistically significant effects (TRANSFORM-2) and one discontinuation trial (SUSTAIN-1).

3.8 MADRS score

We note the following: “The clinical expert noted that MADRS is non-linear, meaning that a change in score at the lower end of the scale does not mean the same, in terms of clinical importance, as a change in score at higher end of the scale.” We are not certain what this claim is based on, but analysis of MADRS and CGI scores demonstrates a linear relationship between MADRS scores and clinical impressions of overall improvement as seen in the figure below from Leucht et al (2017)(Leucht et al., 2017).

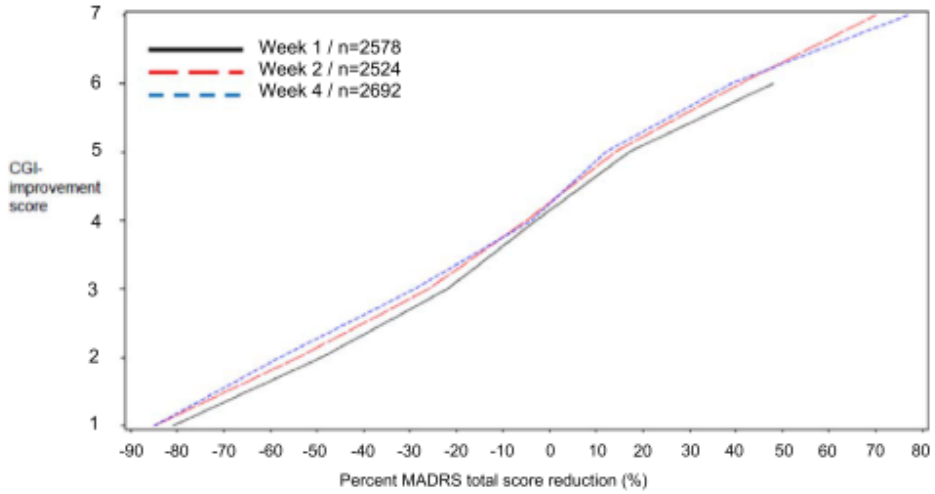


Fig. 2. Linking the percentage change from baseline in the MADRS total scores with the CGI-I scores.

•

As the committee recognises, the definition of 'remission' and 'response', while widely used in academic studies, are by no means standardised or intuitively recognisable entities. Indeed, as mentioned previously the artificial dichotomisation of a continuous scale like the MADRS has the tendency to exaggerate differences between similar groups(Kirsch & Moncrieff, 2007), as appears to have happened in the case of esketamine.

3.9 Time period

Although the committee recognises that 4 weeks trials have little relevance to real world treatment of depression, they comment that further improvements may occur after the 4-week period. This is possible – however, it is more likely that effects will reverse after a period of treatment, as ingestion of psychoactive substances associated with tolerance and withdrawal often do – as in the case of opioids (which have been identified as having some overlap with the mode of action of esketamine)(Schatzberg, 2019) or benzodiazepines, both of which have diminishing effects over time due to tolerance.

We note the BAP has submitted the comment "patients with TRD generally maintained their improvements seen at the end of acute treatment, and even on average improved further," But this statement is not supported by existing evidence. Indeed, the evidence from recreational ketamine users is that long-term use is associated with dysphoria. Increased depression scores were found in both daily users and ex-ketamine users over the course of 1 year(Morgan, Muetzelfeldt, & Curran, 2010), although not in current infrequent users (<3 times per week), so it is possible that this will not apply to esketamine users, but it also seems improbable to predict further improvement based on the common finding of dysphoria in long-term users of this drug (and many other similar substances).

3.10 Power/clinically significant difference.

“The clinical expert commented that for a population in a trial, a mean difference of 4 was clinically significant,” but there is no evidence for this statement. A change in score of 4 or less has been clearly shown to be less than the change required for a clinician to detect a ‘minimally improved’ difference, by the acknowledged leader in the field of establishing minimally clinically significant improvements (Leucht et al., 2017). In this study 22 drug trials, involving 3288 patients were analyses at several time points, with a highly consistent relationship found between MADRS score and CGI-I (an intuitive scale with high inter-rater reliability), finding that the minimally detectable difference by a clinician was 7-9 points. We are not aware of other analyses that have provided analysis on the clinical significance of change in MADRS scores, so it is not clear that it is accurate to say that “there is debate about what is considered a minimal clinically significant difference in the literature.”

Additionally, as the committee has recognised, the studies were powered to find a minimum difference of 6.5 MADRS point between the placebo and esketamine groups, and this means the study may be under-powered to detect a difference as small as 4 MADRS points, indicating the findings may not be reliable.

The response to placebo was indeed large in the esketamine trials. This is perhaps unsurprising given that participants had several hours of contact per week (at least two hours with staff, two times a week, during administration of esketamine and supervision afterwards). Indeed, one interpretation of these trials is that human contact for several hours a week has a large effect on depressed people – a 17 point reduction in MADRS score corresponds to ‘much improved’ (Leucht et al., 2017).

It does not seem reasonable however, to consider that a large placebo response to make it more difficult to ascertain the effects of esketamine: the central premise of a placebo-controlled trial is to identify drug-specific effects; there is no reason to suppose that this was not achieved in this trial.

Regression to the mean is surely a factor in the trial – but again it would affect both arms of the trials equally.

The EMA provided no evidence for considering the effect size to be clinically significant and it is unclear what their opinion was derived from.

3.11 Withdrawal design.

The committee recognises that participants in SUSTAIN-1 were more likely to tolerate the adverse effects of the drug, but additionally they were also selected for inclusion only if they achieved treatment response (>50% reduction in baseline MADRS scores) in the short term trials.(Daly et al., 2018) Therefore the group selected for this trial represent an enriched group of patients who respond positively to the medication and do not represent the wider group of patients who would be given the drug in practice (ie not

already selected for tolerability and response), who would be unlikely to show as large an effect.

As a consequence of choosing ‘responders’, withdrawal designs have been described as tautological because they test whether a treatment shows an effect in a group who have been selected because the treatment shows an effect (that is, it is only ‘responders’ who are recruited into these trials).(Ghaemi & Selker, 2017) It is therefore likely that esketamine would have lesser effects in a less selected population.

As regards unblinding, it is highly implausible, given the immediate physical and mental alterations produced by esketamine (a doubly potent enantiomer of a drug used to produce a ‘high’ for recreational users, at similar doses to that employed in the current trials) that any esketamine trial using a pharmacologically inert placebo could be truly blind. This also applies to withdrawal trials, where participants randomized to placebo will undoubtedly notice that they do not experience the same immediate effects as they experienced before. The FDA analysis of dissociation symptoms (using the CADSS) confirms this (p.28-29 of NDA)(FDA, 2019). The FDA found (in Figure 6 on p.28 of the NDA) that CADSS scores declined rapidly in the arm randomised to placebo. CADSS score was found to significantly associated with time to relapse of depression. The FDA offered several alternative interpretations, the more probable of which was that “the subject may worsen either due to suspecting they are no longer taking active drug.” They also offer another explanation, which is also plausible: that the dissociative effects of esketamine are responsible for slight changes on depression scores because people are literally ‘out of it’, which may reduce the intensity of some domains recorded in the MADRS; when this effect abates, so does the marginal reduction on depression scores.

3.12Withdrawal effects

The point made by the committee in 3.11: “The committee also noted that people with depression in stable response or remission from the TRANSFORM trials who only had placebo had a lower relapse rate than those who stopped esketamine, although this was not explored fully by the company.” The fact that relapse rates were higher in the group discontinued from esketamine than in the untreated group emphasises the fact that removing esketamine is not just revealing the underlying condition (which should produce relapse rates that are equal to the untreated group), but that there is an additional drug withdrawal effect. This suggests that many ‘relapses’ are in fact withdrawal effects, which have been mis-classified as ‘relapse’ because of the overlap between withdrawal effects and domains on the MADRS, as noted by the committee. This interpretation is strongly supported by the pattern of relapses which is visibly seen in the Kaplan-Meier plot reproduced below from a letter by the company in the Lancet Psychiatry(Singh et al., 2020). This figure is notable because it shows any separation between the two arms occurs in the first 4 weeks and that the curves actually cross after 36 weeks with barely any difference between the two. This pattern (large differences at short time periods, converging at long

time periods) is the hallmark of a withdrawal syndrome: withdrawal symptoms cluster towards the point of cessation and slowly resolve over weeks or months.

“The company considered that there would be no long-term withdrawal effects of esketamine because at this dose it leaves the body quickly.” This is inaccurate. Withdrawal effects arise because the body has adapted (become tolerant) to the presence of the drug (perhaps up-regulating NMDA receptors in the case of an NDMA antagonist, although there are many potential neurobiological mechanisms of tolerance/adaptation to esketamine) which is then removed. Withdrawal effects are the subjective experience of a system which has become homeostatically ‘tuned’ to a certain level of drug then being subjected to lower levels of it. The more quickly a drug leaves the body the more quickly and severe will be the withdrawal effects. For example, short acting benzodiazepines, antidepressants with short half-lives (eg paroxetine) and opioids with short half-lives produce the most severe withdrawal symptoms. *Withdrawal symptoms persist for the period of time taken for adaptations in the brain and body to resolve back to a pre-drug state – not for the period of time taken for a drug to leave the body.* This process of resolving adaptation can take weeks or months (although possibly less after short-term use). It is therefore misleading to suggest that a short half-life for a drug is consistent with there being no withdrawal effects; it is the opposite that is true.

The clinical expert is misleading when he states: “that withdrawal effects of ketamine seen in recreational use are from higher doses.” In fact, doses of esketamine employed in the trials are similar to those used recreationally. Esketamine doses employed by Janssen (56-84mg), were not distinct from those used by recreational users (equivalent to 50-100mg esketamine),(Sassano-Higgins, Baron, Juarez, Esmaili, & Gold, 2016) noting that esketamine has twice the potency of ketamine. Moreover, we know from experience with other drugs that, there is no threshold for which withdrawal symptoms occur and no reason to think they would not occur at the doses of esketamine employed by Janssen. Although ‘sweating and shaking’ may be more common with withdrawal from higher doses, withdrawal commonly involves numerous other symptoms such as fatigue, poor appetite, drowsiness, anxiety and dysphoria, which, as the committee note, overlap with items on the MADRS and may be mistaken for a relapse of depression.(Chen, Huang, & Lin, 2014; Cosci & Chouinard, 2020; Morgan & Curran, 2012)

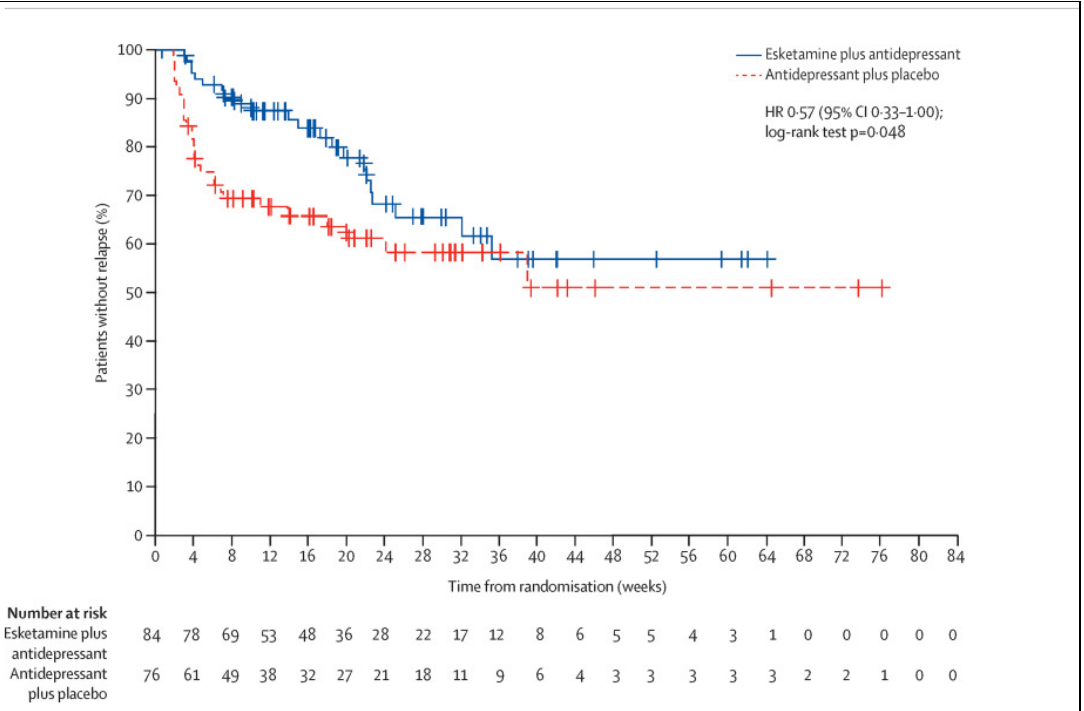


Figure Kaplan-Meier estimates of time to relapse among patients who achieved stable remission

Hide caption

3.13Outlier site.

We maintain that it is reasonable to exclude the outlier site and we point out that the EMA did not provide any rationale for retaining it. The data is highly inconsistent with the data from other sites. Analysis of this study without this site showed there was no significant difference between the two arms, using Fisher's exact test (p=0.13).(Turner, 2020) This test is more appropriate than the log-rank method preferred by the drug manufacturer, because the log-rank method exaggerates the effect of early relapses, inappropriate in this case where these relapses are most likely to reflect withdrawal-related events.(Singh et al., 2020) The sponsor's re-analysis, using the log-rank method, excluding the outlier site yielded a barely significant P=0.048, but one wonders whether the FDA's analysis maintained significance, as they stated several times in the regulatory review documents that the outlier site "drives" the overall study results.(Turner, 2020) Although patient-level data was requested in order to resolve this issue, the manufacturer has not, to our knowledge, furnished independent researchers access to this.

3.14Exclusions.

The manufacturer excluded patients who would meet the usual clinical definition of 'severe depression', including suicidal ideation (in the last 6 months), suicide attempts (in the last 12 months), co-morbidity (drug and alcohol and mental conditions) and past history of ECT. It would seem that the patient population was a much less severe group of patients than clinicians would generally consider to be 'severely depressed' or 'treatment resistant'. This notion is emphasised by the fact that these 'treatment

resistant' patients showed a large response (17 MADRS points) to placebo (human contact). This makes the lack of clinical efficacy and the numerous safety signals for esketamine even more concerning – and the fact that a number of these patients committed suicide, despite having been selected for lack of suicidality.

3.15Placebo adjustment

The suggestion by the manufacturer that the placebo arm of a placebo-controlled trials is adjusted down violates the basic premise of placebo-controlled trials. Every point raised by the company as to why placebo response rates were high also relate to the esketamine arm.

3.16Safety

A salient point was made in the BMJ this month by John Warren, former Expert Medical Assessor and NDA evaluator for the MHRA: that in trials, manufacturers use composite scores in order to find positive effects for their drugs but use separate events for side effects, which has the effect of minimising the overall burden of negative effects.(Warren, 2020) This was evident in the esketamine trials, which was identified as a case in point. MADRS is a composite score which measures 10 symptom domains, including appetite changes, mood, sleep etc. Composite measures are more likely to find differences between groups (as small effects add up). In contrast, the many 'side effects' of esketamine were grouped individually. As stated in the BMJ article: "An incidence of at least 5% and at least twice that of placebo was reported for dissociation, dizziness, nausea, sedation, vertigo, hypoaesthesia, anxiety, lethargy, blood pressure increase, vomiting, and feeling drunk.(8) Whereas one primary endpoint was used to summarise benefit, safety was analysed as a collection of symptoms with no single endpoint, mitigating against finding statistical significance (32) and leading to the asymmetrical analysis of risk and benefit.(33)"

Post-marketing surveillance of esketamine has only served to strengthen these concerns with strong signals emerging for the data for an increased risk of dissociation, sedation, feeling drunk, suicidal ideation and completed suicide(Gastaldon, Raschi, Kane, Barbui, & Schoretsanitis, 2020). Although such post-marketing surveillance data are inherently limited in drawing conclusions about causality, because they are not randomised participants, confounding by indication is possible, and there may be a notoriety bias in reporting, many of these effects (including suicidal ideation) remained when comparisons were made to adverse effects reported for venlafaxine and, furthermore, the effects reported closely mirror those reported in the regulatory trials submitted to the FDA, triangulating the evidence.

The increased rates of worsening of depression and suicidal ideation in the esketamine group compared with the placebo group in the regulatory trials, although small in number, is another signal consistent with the risk for deterioration and increased suicidality in esketamine use. This may well be

explained by the intense dysphoria that some users can experience during and after short-term treatment (see patient extract below in Appendices) and which is reported to occur with long-term use among recreational users (Morgan & Curran, 2012)

One important issue not as remarked upon in the otherwise comprehensive appraisal is the strong association of motor vehicle accidents with esketamine use. Esketamine is known to impair hand-eye coordination, to cause dissociation, and to be associated with car accidents(Cheng, Chan, & Mok, 2005; Morgan & Curran, 2012; Schifano, Corkery, Oyefeso, Tonia, & Ghodse, 2008). In the regulatory trials there were 5 car accidents all in the esketamine arm, one of which was fatal. As can be seen in the supplementary material of Gastaldon et al. (2020) a number of road traffic accidents were attributed to esketamine in the FDA database. In Hong Kong where recreational ketamine achieved particular popularity in the 1990s 9% of all fatal traffic accidents from 1996 to 2000 involved ketamine use (Cheng et al., 2005).

There are also some elements of the safety study design (SUSTAIN-2) (Wajs et al., 2020)that do not lend it to a full understanding of the adverse effects of esketamine. 802 patients were enrolled into this study but 331 were excluded as the ‘study was stopped by sponsor’. The explanation given for excluding almost half the patients was unclear. In the remaining patients there were some concerning findings: 114 patients had new onset suicidal ideation (in a group selected for lacking any suicidality), there were 6 suicide attempts and one completed suicide (in a patient with no history of suicidal behaviour or intent).

There is some speculation that the increased suicidality seen in esketamine use may be related to its psychedelic properties (which may be useful for some, but lead to terror and fear for others). The dissociation caused by esketamine may be one manner of describing the hallucinogenic and psychedelic properties of the drugs.

Lastly, there is concern that NMDA antagonists can be neurotoxic in the long term. Ketamine was originally developed from phenylcyclidine (PCP), known as ‘angel dust’ when used recreationally. Ketamine and PCP have similar chemical structures, and are both primarily NMDA antagonists although the potency of PCP is greater than ketamine. PCP causes similar effects to ketamine: hallucinations, distorted perception of sound (see patient accounts below). It also causes an increased risk of suicide, which some have linked to its ability to produce flashbacks (and some have linked ketamine suicides to flashbacks as well). It is also used as an anaesthetic agent and can cause euphoria in the short-term. It is, like ketamine, psychotomimetic (Murrie, Lappin, Large, & Sara, 2020). Both PCP and ketamine are highly lipid soluble, and ketamine possesses a chlorine atom (halogenation is widely used in anaesthetics to enhance penetration of the brain as it dramatically increases transfer of the molecule across cell membranes).

The dangers of NMDA antagonists like ketamine have been demonstrated that repeated exposure to ketamine-like drugs during development can permanently disrupt neurodevelopment and have catastrophic long-term cognitive and behavioural outcomes (Zimmermann, Richardson, & Baker, 2020). In animal models, exposure to NMDA receptor antagonists during development in animal models impairs parvalbumin maturation, reduces the number of parvalbumin neurons in the medial prefrontal cortex and causes disorganised prefrontal cortex output in adulthood, mimicking the disease pathology of schizophrenia (Zimmermann et al., 2020). Notably, the doses used to induce schizophrenia-like dissociative symptoms and disrupt parvalbumin development in animals are similar to the doses used to in the Janssen depression trials (Zimmermann et al., 2020).This is reflected in its pro-psychotic effects in adult subjects given doses of esketamine used in the depression trials (Beck et al., 2020), although possibly the neurotoxic effects of esketamine might be less potent in adult brains. However, this possibility has not been excluded with long-term safety trials focusing particularly on brain effects. The cognitive impairment and dysphoria seen in recreational ketamine users is not reassuring (Morgan & Curran, 2012).

Placebo-controlled long-term studies of safety utilising composite scales to assess for a range of safety side effects, with particular focus on neural effects of long-term administration of esketamine are required for the medication is released into general use. A Ketamine Side Effect Tool and Ketamine Safety Screening Tool have been developed by ketamine experts in Australia for this purpose(Short, Fong, Galvez, Shelker, & Loo, 2018). 13 major side effects are included based on a systematic review of the literature, including headache, dizziness, dissociation, increased blood pressures, blurred vision, nausea, sedation or drowsiness, faintness or light-headedness, anxiety, elevated heart rate, cognition side effects, urinary tract side effects and dependency risk (Short et al., 2018). Although severity ratings would be required to adequately match the level of detail captured in composite measures of efficacy such as the MADRS, an initial assessment of data collected by Janssen using this composite scale would be instructive to compare with the positive effects of the drug.

3.22 Adjustment to mortality

It was also concerning to see the manufacturer attempt to suggest that mortality would be improved, when there were more deaths in the esketamine group than the placebo group (even taking into account that patients spent longer on esketamine than placebo). There was barely any discernible effect of esketamine on depression scores to justify the calculation they proposed, and projection over time of results found in 4 week trials is unreasonable especially given the experience in recreational use with ketamine that it is associated with dysphoria in the long-term (Morgan & Curran, 2012).

3.25 Stopping treatment

As with all drugs that cause dependence and withdrawal, and can cause addiction, it will likely be hard for people to stop esketamine after a period of use, which will contribute to long-term use. Some of these patients might resort to buying ketamine off the streets as occurred with opioids and heroin in the US – despite the view offered by the manufacturer that addiction is not possible because use is supervised. Addicted people will find a means to obtain a supply.

As has happened with benzodiazepines, and now antidepressants, people who stop their esketamine are likely to experience withdrawal symptoms and as the manufacturer is informing doctors and patients that withdrawal effects are not possible, patients will be diagnosed with 'relapse' of their condition, and will be told to re-start treatment. In this manner many people will be caught in long-term use by withdrawal symptoms – a very common phenomenon in psychiatry as highlighted by the recent PHE report on prescribed drug dependence(Public Health England, 2019) in which they said: "Recurring patterns are evident in the history of medicines that may cause dependence or withdrawal. New medicines are seen as an important part of the solution to a condition, resulting in widespread use. Their dependence or withdrawal potential are either unknown at this point, due to a lack of research, or perhaps downplayed. As evidence of harm from dependence or withdrawal emerges, efforts are made to curtail prescribing. The repetition of this pattern is striking."

3.33 Innovative action

'Innovative action' is not a positive thing in its own right in the absence of evidence for meaningful efficacy and safety. This is a marketing device, not a serious point. It is like saying 'we have installed new breaks on your car: they are a totally new design! But they don't work very well and sometimes catch on fire.'

In relation to the point "esketamine is sprayed in the nose which means it works rapidly": depression is a chronic condition and the ability to quickly absorb something through the nose is of no consequence in the time scale that matters. Other anaesthetic agents such as propofol (Mickey et al., 2018) and nitrous oxide (Nagele et al., 2015) also reduce depression scores in a few hours, but it seems improbable that this represents a sustainable effect on a long-term condition and is likely broadly similar to the effects of esketamine.

Kind regards,

[redacted], Clinical Research Fellow, [redacted], Professor of Critical and Social Psychiatry, [redacted], Oregon Health and Science University [redacted], Private Practice, [redacted], Stockholm, Sweden.

Appendix

Patients have reported their anecdotal experiences of using esketamine online: https://www.drugs.com/comments/esketamine/spravato.html

Some people report positive effects, some negative. There are also reports like this which are consistent with a dysphoric effect from a psychedelic and might explain the potential for increased suicidal ideas and attempts amongst some users:

“On Spravato I felt like I was going to die. It took 4 minutes to kick in. I couldn’t feel my face, couldn’t speak, couldn’t breathe, the room was spinning, music felt far away even though I had headphones in. I heard the ocean. I was sweating/ felt like I was sinking, mind racing, nauseous. I felt like I was strapped to a Ferris wheel that wouldn’t stop spinning/falling really fast. I was crying & sweating, 22 minutes in I had a ringing in my ear. It gave me a headache. I couldn’t hear my music anymore, that scared me. I couldn’t move. I was crying for my mom. It was really scary. I don’t feel 100% back to normal today (appointment was yesterday). I feel lethargic-spacey with no appetite. I feel sad/disappointed that this didn’t work. This was literally the last thing to try treatment wise. I have to go back to doing TMS everyday. Please don’t try this unless it’s your LAST resort.”

Here are some further reports. These represent people who had a negative experience and did not feel it was beneficial, and it should be noted there are also a number of more positive experiences where people report an improvement in their depressive symptoms.

“I suffer from depression. I’ve tried different meds and none have seemed to help. I heard of Spravato and was very excited to know there was a treatment out there that could work. I’ve done 6 sessions of Spravato and have seen no improvement. I have an appointment tomorrow for another treatment. I really don’t like the experience. As you inhale the Spravato you get the nasty taste. Then you’re in for a ride of feeling like you’re outside your body. Feeling like you’re completely numb. Face feet etc. I hate the whole experience. Everyone is different. I hope this works for others that are going to try this. Good luck!”

“Anybody have increased suicidal thought on spravato?”

“I have Bipolar, untreatable depression, BPD. I have done esketamine as both IV and Spravato. Spravato is a waste of time and money. After going through the initial 8 week transition, I was in for Spravato almost every other week where I was in for IV once a month to as long as every 6 weeks. The side effects were horrible, headache, exhaustion for a day or 2 and not knowing where I was. This drug was barely vetted by the FDA with only 80 test subjects in their group. IV ketamine has years of use and they know it

works, but because there is zero money to be made they will never approve it so our insurance will pay for it”

“After being on 6 different anti depressants and doing 6 weeks of TMS with no relief from my depression I was prescribed Spravato. I completed 4 weeks of treatments of 84mg 2 times per week, 4 weeks of treatments one times per week and one treatment every other week. I have had no relief from my depression. I feel like the effectiveness of this drug was over hyped. It was expensive and time consuming and very disappointing to have no results.”

“I had my first experience with Spravato yesterday and the side effects were just awful. Felt extremely drunk--dissociation, severe nausea and sweating, dizziness, tiredness. Could barely make it home (I got a ride) and get myself into bed to lie there for the rest of the night. The next morning--extremely tired, nauseous. Felt hungover. No affect on my TRD (treatment resistant depression after 1 treatment. I will continue as long as I can stand it because I am pretty much out of options. But I am dreading it more than I did with ECT.”

“I posted my experience earlier and now I have done a total of 10 treatments of the nasal spray Spravato. 2 at 56mg and the rest at 84mg. After sessions 6 I started to feel a little lighter and less depressed but this lasted very shortly. The treatments after that did nothing for me. They are very intense and you get awful disassociation, dizziness, numbness. Its a complete out of body experience. For me the feelings were very bothersome & after I would leave I was done for the day. I couldn’t do anything just lay around. I guess if it worked for me I would have sucked it up & continued even though the side effects were so bothersome. Unfortunately for me this didn’t work and it cost me a fortune. Personally I wouldn’t advise others to do this until there is more research about it. It is still very new and drs are still learning by experimenting on us! However the one done by IV is not as new. I will continue to fight the fight but will never do Spravato again.”

“So far I am on my 3rd time of doing the Spravato Nasal spray. 2x I did 56mg and one time the 82mg I believe. I feel exactly the same. No change at all!! My Dr told me to at least give it to the one month mark. It makes me feel awful during it I get the disassociation and dizziness along with every other side effect and I will say it takes me more than 2 hours to wear off. Pretty much the whole day. I am hoping the next few treatments work otherwise I have. Wasted a lot of time and money:-(”

For Depression: “TRD (treatment resistant depression) for 30 years, have tried everything and will continue to try this. Was on 56mg 2x week for a month (eventually I had no immediate reaction as my body got used to the drug - lingering tiredness, so I did the treatments in the late afternoon) and have had 3 treatments with 84mg. First two of 84mg were fine - trippy, slight visual disturbance but I was able to recognize that this was the drug and not me - felt like I smoked a ton of pot or something. Yesterday was the 3rd treatment of 84mg. I went into appointment very upset and stressed, and

coupled with the Spravato, had a dissociative episode. I started to cry and then boom: sobbing, no idea who I was, where I was, visual disturbance so couldn't see correctly, was paranoid, untrusting - I have never been so terrified in all my life. So far, no depression reduction, which saddens me, and I'm scared to have another trip like I did yesterday."

There has been a paper published looking at the adverse events from esketamine reported to the FDA for the 12 months that it has been in use there that should be very pertinent to the committee's deliberations, attached, and which I hope there will be scope to consider.

Comparing adverse events reported for esketamine to all other drugs in the FDA's adverse events reporting system, they found for esketamine a 1600-fold increase in dissociation, a 240-fold increase in sedation, 100-fold increase in 'feeling drunk', as well as a 24-fold increase in suicidal ideation and a 6-fold increase in completed suicides, compared with other drugs. Although confounding by indication is possible, esketamine is offered to not particularly suicidal people (at least 44% of patients will fail two antidepressants from the STAR-D trial) and signals regarding suicide remained even when comparing to venlafaxine. The risks found reflect the findings of the regulatory studies for esketamine where all of these signals were found (but minimised).

The authors recommend 'urgent clarification' of the suicidal effects of the drug.

I hope that the NICE committee will take this into account when considering approving this drug for use in the UK.

References

Beck, K., Hindley, G., Borgan, F., Ginestet, C., McCutcheon, R., Brugger, S., … Howes, O. D. (2020). Association of Ketamine With Psychiatric Symptoms and Implications for Its Therapeutic Use and for Understanding Schizophrenia: A Systematic Review and Meta-analysis. *JAMA Network Open*, *3*(5), e204693–e204693. https://doi.org/10.1001/jamanetworkopen.2020.4693

Canuso, C. M., Singh, J. B., Fedgchin, M., Alphs, L., Lane, R., Lim, P., … Drevets, W. C. (2018). Efficacy and safety of intranasal esketamine for the rapid reduction of symptoms of depression and suicidality in patients at imminent risk for suicide: Results of a double-blind, randomized, placebo-controlled study. *American Journal of Psychiatry*, *175*(7), 620–630. https://doi.org/10.1176/appi.ajp.2018.17060720

Chen, W. Y., Huang, M. C., & Lin, S. K. (2014). Gender differences in subjective discontinuation symptoms associated with ketamine use. *Substance Abuse: Treatment, Prevention, and Policy*, *9*(1), 1–7. https://doi.org/10.1186/1747-597X-9-39

Cheng, J. Y. K., Chan, D. T. W., & Mok, V. K. K. (2005). An epidemiological study on alcohol/drugs related fatal traffic crash cases of deceased drivers in Hong Kong between 1996 and 2000. *Forensic Science*

International, *153*(2–3), 196–201. https://doi.org/10.1016/j.forsciint.2004.08.023

Cosci, F., & Chouinard, G. (2020). Acute and Persistent Withdrawal Syndromes Following Discontinuation of Psychotropic Medications. *Psychotherapy and Psychosomatics*, *89*(5), 283–306. https://doi.org/10.1159/000506868

Daly, E. J., Singh, J. B., Fedgchin, M., Cooper, K., Lim, P., Shelton, R. C., … Drevets, W. C. (2018). Efficacy and safety of intranasal esketamine adjunctive to oral antidepressant therapy in treatment-resistant depression: A randomized clinical trial. *JAMA Psychiatry*, *75*(2), 139–148. https://doi.org/10.1001/jamapsychiatry.2017.3739

Fabbri, A., Lai, A., Grundy, Q., & Bero, L. A. (2018). The Influence of Industry Sponsorship on the Research Agenda: A Scoping Review. *American Journal of Public Health*, *108*(11), e9–e16. https://doi.org/10.2105/AJPH.2018.304677

Fabbri, A., Parker, L., Colombo, C., Mosconi, P., Barbara, G., Frattaruolo, M. P., … Mintzes, B. (2020). Industry funding of patient and health consumer organisations: systematic review with meta-analysis. *BMJ*, *368*, l6925. https://doi.org/10.1136/bmj.l6925

FDA. (2019). Efficacy, safety, and risk-benefit profile of New Drug Application (NDA) 211243, esketamine 28 mg single-use nasal spray device, submitted by Janssen Pharmaceuticals, Inc., for the treatment of treatment-resistant depression, 1–135.

Fu, D. J., Canuso, C. M., Ionescu, D., Li, X., Lane, R., Lim, P., … Manji, H. (2019). *Esketamine nasal spray for rapid reduction of major depressive disorder symptoms in patients at imminent risk for suicide: ASPIRE-1 study. Poster presented at the Psych Congress, October 3-6, 2019, San Diego, CA*. Retrieved from https://www.psychcongress.com/sites/psychcongress.com/files/2019-12/138-Esketamine-nasal-spray-for-rapid-reduction-of-.pdf

Gastaldon, C., Raschi, E., Kane, J. M., Barbui, C., & Schoretsanitis, G. (2020). Post-Marketing Safety Concerns with Esketamine: A Disproportionality Analysis of Spontaneous Reports Submitted to the FDA Adverse Event Reporting System. *Psychotherapy and Psychosomatics*, 1–8. https://doi.org/10.1159/000510703

Ghaemi, S. N., & Selker, H. P. (2017). Maintenance efficacy designs in psychiatry: Randomized discontinuation trials – enriched but not better. *Journal of Clinical and Translational Science*, *1*(3), 198–204. https://doi.org/10.1017/cts.2017.2

Ionescu, D. F., Canuso, C. M., Fu, D. J., Qiu, X., Lane, R., Lim, P., … Manji, H. (2019). *Esketamine nasal spray for rapid reduction of major depressive disorder symptoms in patients at imminent risk for suicide: ASPIRE-2 study. Poster presented at the: Psych Congress, October 3-6, 2019, San Diego, CA*. Retrieved from https://www.psychcongress.com/sites/psychcongress.com/files/2019-12/151-Esketamine-nasal-spray-for-rapid-reduction-of-.pdf

Jakobsen, J. C., Katakam, K. K., Schou, A., Hellmuth, S. G., Stallknecht, S. E., Leth-Møller, K., … Gluud, C. (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis.

BMC Psychiatry, *17*(1), 58. https://doi.org/10.1186/s12888-016-1173-2
Kirsch, I., & Moncrieff, J. (2007). Clinical trials and the response rate illusion. *Contemporary Clinical Trials*, *28*(4), 348–351. https://doi.org/https://doi.org/10.1016/j.cct.2006.10.012
Leucht, S., Fennema, H., Engel, R. R., Kaspers-Janssen, M., Lepping, P., & Szegedi, A. (2017). What does the MADRS mean? Equipercentile linking with the CGI using a company database of mirtazapine studies. *Journal of Affective Disorders*, *210*(September 2016), 287–293. https://doi.org/10.1016/j.jad.2016.12.041
Lock, I., Seele, P., & Heath, R. L. (2016). Where Grass Has No Roots: The Concept of 'Shared Strategic Communication' as an Answer to Unethical Astroturf Lobbying. *International Journal of Strategic Communication*, *10*(2), 87–100. https://doi.org/10.1080/1553118X.2015.1116002
Mickey, B. J., White, A. T., Arp, A. M., Leonardi, K., Torres, M. M., Larson, A. L., … Tadler, S. C. (2018). Propofol for treatment-resistant depression: A pilot study. *International Journal of Neuropsychopharmacology*, *21*(12), 1079–1089. https://doi.org/10.1093/ijnp/pyy085
Morgan, C. J. A., & Curran, H. V. (2012). Ketamine use: A review. *Addiction*, *107*(1), 27–38. https://doi.org/10.1111/j.1360-0443.2011.03576.x
Morgan, C. J. A., Muetzelfeldt, L., & Curran, H. V. (2010). Consequences of chronic ketamine self-administration upon neurocognitive function and psychological wellbeing: A 1-year longitudinal study. *Addiction*, *105*(1), 121–133. https://doi.org/10.1111/j.1360-0443.2009.02761.x
Murrie, B., Lappin, J., Large, M., & Sara, G. (2020). Transition of Substance-Induced, Brief, and Atypical Psychoses to Schizophrenia: A Systematic Review and Meta-analysis. *Schizophrenia Bulletin*, *46*(3), 505–516. https://doi.org/10.1093/schbul/sbz102
Nagele, P., Duma, A., Kopec, M., Gebara, M. A., Parsoei, A., Walker, M., … Conway, C. R. (2015). Nitrous oxide for treatment-resistant major depression: A proof-of-concept trial. *Biological Psychiatry*, *78*(1), 10–18. https://doi.org/10.1016/j.biopsych.2014.11.016
Public Health England. (2019). Dependence and withdrawal associated with some prescribed medicines. An evidence review. Retrieved from https://www.gov.uk/government/publications/prescribed-medicines-review-report
Ruhé, H. G., van Rooijen, G., Spijker, J., Peeters, F. P. M. L., & Schene, A. H. (2012). Staging methods for treatment resistant depression. A systematic review. *Journal of Affective Disorders*, *137*(1), 35–45. https://doi.org/https://doi.org/10.1016/j.jad.2011.02.020
Sassano-Higgins, S., Baron, D., Juarez, G., Esmaili, N., & Gold, M. (2016). A Review of Ketamine Abuse and Diversion. *Depression and Anxiety*, *33*(8), 718–727. https://doi.org/10.1002/da.22536
Schatzberg, A. F. (2019). A word to the wise about intranasal esketamine. *American Journal of Psychiatry*, *176*(6), 422–424. https://doi.org/10.1176/appi.ajp.2019.19040423
Schifano, F., Corkery, J., Oyefeso, A., Tonia, T., & Ghodse, A. H. (2008). Trapped in the "K-hole": Overview of Deaths Associated With Ketamine Misuse in the UK (1993-2006). *Journal of Clinical Psychopharmacology*,

28(1).
Short, B., Fong, J., Galvez, V., Shelker, W., & Loo, C. K. (2018). Side-effects associated with ketamine use in depression: a systematic review. *The Lancet. Psychiatry*, *5*(1), 65–78. https://doi.org/10.1016/S2215-0366(17)30272-9
Singh, J. B., Daly, E. J., Mathews, M., Fedgchin, M., Popova, V., Hough, D., & Drevets, W. C. (2020). Approval of esketamine for treatment-resistant depression. *The Lancet Psychiatry*, *7*(3), 232–235. https://doi.org/10.1016/S2215-0366(19)30533-4
Turner, E. H. (2020). Approval of esketamine for treatment-resistant depression – Author's reply. *The Lancet Psychiatry*, *7*(3), 236. https://doi.org/10.1016/S2215-0366(20)30044-4
Wajs, E., Aluisio, L., Holder, R., Daly, E. J., Lane, R., Lim, P., … Singh, J. B. (2020). Esketamine Nasal Spray Plus Oral Antidepressant in Patients With Treatment-Resistant Depression. *The Journal of Clinical Psychiatry*, *81*(3), 4–5. https://doi.org/10.4088/JCP.19m12891
Warren, J. B. (2020). The trouble with antidepressants : why the evidence overplays benefits and underplays risks — an essay by John B Warren. https://doi.org/10.1136/bmj.m3200
Zimmermann, K. S., Richardson, R., & Baker, K. D. (2020). Esketamine as a treatment for paediatric depression: questions of safety and efficacy. *The Lancet Psychiatry*, *0366*(19), 10–11. https://doi.org/10.1016/s2215-0366(19)30521-8

Name	[redacted]
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
I'm afraid I do not think so. As I don't think the evidence related to the direct and indirect cost of depression has been fully appreciated by the committee.

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
No.

There is a huge unmet need for patients who fail to respond to traditional antidepressants. According to the available evidence I believe that Esketamine is cost-effective if prescribed to the right group of patients.

Are the recommendations sound and a suitable basis for guidance to the NHS?
No

Name	[redacted], [redacted]

Comments on the ACD:

Has all of the relevant evidence been taken into account?
Yes

General comment
As a consultant psychiatrist I write to recommend that this medicine be given approval in the England and Wales and Northern Ireland, following the example set by the Scottish Medicines Consortium earlier this year.

I have worked for many years with patients suffering from therapy resistant depression, which is a totally debilitating condition, meaning that otherwise intelligent people are required to be signed off from their workplaces for years at a time.

One of our patients has been part of Phase 3 clinical trial on Esketamine nasal spray. The effect of this trial has been life changing and has completely reversed the disorder, and our patient is now able to benefit from the talking therapy she has been undertaking throughout her disorder, as she can now feel the benefits of the work she has been doing on her mood, rather than just intellectually understanding what her therapist has been discussing with her, but it having no effect on her mood.

I respectfully submit to the committee that this medication should be approved for use in England Wales and Northern Ireland for the benefit of those receiving treatment on the NHS.

Yours faithfully

[redacted]
MB ChB, MD, PhD, Swedish Board of Psychiatry
Consultant Psychiatrist
GMC no [redacted]

Name	[redacted]

Comments on the ACD:
General comment I feel that with Covid being a likely ongoing issue, that this may be an alternative to ECT. ECT could not be offered in a Covid safe way, but this may offer a realistic alternative using the same staffing and resources.

Name	[redacted], College of Mental Health Pharmacy
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
p13. 3.10 It seems unfair to criticise the trial design when the design was dictated for licensing by the MHRA and FDA. The FDA/MHRA did not allow a placebo group because TRD is such a serious condition that it would be considered unethical to allocate some people to a placebo.
Treatment-Resistant Depression is a condition people live with every day and is so severe, horrible and painful that some find suicide the only option. https://www.theguardian.com/society/2020/sep/01/male-suicide-rate-england-wales-covid-19

There are very few other conditions where the symptoms are so unpleasant people commit suicide because of them. And many who do not end their lives only do so because of the effect that would have on their families.
As you acknowledge, hopelessness is a predominant symptom of TRD and, quite often, it is a realistic and comprehendible hopelessness.

page 13 3.11: trial design: We appreciate that it is right that the ERG will be concerned about adverse effects. However someone with TRD may be willing to tolerate a much higher side effect burden because the symptoms of TRD are so deeply unpleasant. Moderate side effects are a long way from the actual symptoms of TRD so please don't assume side effects will put people off. When you're in a hopeless worthless state transient side effects (as opposed to continued medicine taking) can be tolerated for the final relief.

page 9 3.6 :
We do not know where the ERG has got the idea that it is a problem that someone couldn't have CBT immediately after esketamine. We are not aware that anyone suggested it could be.
ECT would have exactly the same "restriction".
A distinct advantage of esketamine is that, while it produces side effects at the time, these are limited to the day of treatment which then leaves the person side effect-free on the other 5 or 6 days. This contrasts with oral antidepressants where side effects are continuous and may impede CBT's effectiveness.
We would be fascinated to read a NICE appraisal of CBT with the same degree of forensic analysis of studies as it would be unlikely to prove CBT is

cost effective. We feel it unlikely patients would even want CBT on a day they have had other treatments.

Trends Psychiatry Psychother 2020;42(1):92-101.
Cognitive-behavioral therapy for treatment-resistant depression in adults and adolescents: a systematic review
Stephanie Zakhour 1 , Antonio E Nardi 1 , Michelle Levitan 1 , Jose Carlos Appolinario 1 2 3 4 PMID: 32130308

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
Page 3: See comment in section 1 regarding the FDA and MHRA licensing requirements. The analysis should identify that the licensing trials required by the FDA and MHRA led to them awarding a licence and that they did not see this as compromising proof of efficacy.

It is true that people may be more unwell than in the clinical trials. This is true of almost every condition with a new therapy and TRD is no different. We are unclear on the significance of this statement

Alternatives: Psychological therapies seem to get a disproportionate degree of prominence in the TA.

It is worth knowing that there are few trials of CBT in TRD and it is inappropriate to mention CBT as if to suggest it were a comparable therapy. The most recent systematic review (Zakhour et al, 2020) identified 8 studies of CBT in TRD, although one was an open trial and one was a case report. The remaining six were Randomised Controlled Trials but used waiting list controls (as opposed to double-blind placebo-controlled studies; waiting list controls are where the intervention is compared with someone having no treatment at all, told they were ill and could be treated but have to go on a 16 week waiting list). You cannot account for the impact of being on the waiting list group (which NICE itself pointed out was worse than a placebo). These CBT trial designs are something that, as general medicine clinicians, you would not find acceptable as proof of efficacy in any general medical drug. Thus, although CBT is an option and clearly helps some people, it has little robustly proven efficacy in TRD, or none if you do not accept waiting lists as an adequate control group.

page 3 point 1: there is no evidence that we are aware of that supports this statement. It would seem logically virtually inconceivable that improving TRD will NOT improve someone's quality of life. We feel that is a distracting uncertainty. The ERG even acknowledges the wider effects of TRD p5, 3.1: "The committee concluded that the condition has a negative effect on people, their families and carers."
If TRD has that impact how could going into remission not improve QoL? We would request the ERG re-considers this statement, the evidence supporting it and that this distracts the reader.

page 18 3.16: We think that the ERG has greatly over-emphasised the risks of diversion and abuse. Esketamine will, for many years, be managed by MH Trusts. It is a schedule 2 Controlled Drug, with records and stock control of purchase, supply, administration and destruction. The risk of diversion of a medicine in a countable pack (as opposed to e.g. a liquid in a stock bottle) is close to zero and would be rapidly identified even if only one single pack disappeared. To imply that MH Trusts could not manage this is incorrect.

For someone to abuse esketamine you would first have to prove you had TRD to a Consultant Psychiatrist. You would have a first dose, be monitored, the dose stays stable and then reduces (unlike ketamine, where the dose is increased in order to reach the same effect), you can't take the medicine away with you, it is given under direct supervision, and ketamine is readily available on the streets. On discussion with a DrugScience colleague they felt abuse of esketamine under the strict management conditions highlighted above would be very unlikely with ketamine readily available at £25-30 per gram on the streets.

page 6 3.3: It would be almost inconceivable that someone with TRD would have not had at least two of these antidepressants so the option list looks rather more impressive than it is. They all have similar modes of action on serotonin and/or noradrenaline, and STAR*D shows that response decreases significantly as the number of antidepressants tried increases.

page 11 3.9: We are surprised at how negatively this has been interpreted. A separation from placebo after 2 days in TRD? why is this not considered a true effect?

A 21-point decrease in MADRS scores cf. baseline is, to us , not open to uncertainty.

page 11-12, 3.9: we would challenge the 4 week comment in relation to patient experience. Going into remission within 4 weeks would be incredibly helpful to many patients with TRD.

page 14, 3.12: withdrawal effects: We are pleased that the ERG accepted this. It is important that ketamine is seen as different in this respect, due to escalating doses used by people abusing ketamine and reducing doses for esketamine in depression (after the first dose). It is worth recalling that esketamine is already available on the UK market as an anaesthetic.

P14-15 3.12: withdrawal effects would be difficult to distinguish. we do not support this statement and would be surprised if this was agreed with by those with mental health experience. clinically you would not expect to see long withdrawal symptoms lasting days and looking like depression from something with a relatively short half-life where the central effect is virtually gone from the body in a few hours and which is only taken once or twice a week, not continuously. Any such withdrawal symptoms would not in practice be confused with a relapse of depression. Withdrawal symptoms

would change and reduce over the time between doses. Depression would not improve and would most likely get worse.

page 18 3.15: We do not know why the ERG has got this idea that it is a problem that someone couldn't have CBT immediately after esketamine. We are not aware that anyone suggested it could be.
ECT would have exactly the same "restriction".
A distinct advantage of esketamine is that, while it produces side effects at the time, these are limited to the day of treatment (and essentially only for a couple of hours), which then leaves the person side effect-free on the other 5 or 6 days and more able to make the most of CBT. This contrasts with oral antidepressants where side effects are continuous and may impede CBT's effectiveness.

P18 3.16: the issue of abuse and diversion again: There may be an increased risk but, as predominantly secondary care pharmacists, we do not recognise this as a risk that cannot be managed. Widespread diversion is hard to imagine and there are plenty of other drugs with higher abuse potential and higher availability. Schedule 2 drugs are widely used e.g. ketamine, diamorphine etc and whatever risk there is can be managed within existing systems. Use of standard Controlled Drug systems (especially with a single-use mechanism in distinct boxes) eliminate this risk. As a schedule 2 Controlled Drug, there are records and stock control of purchase, supply, administration and destruction. as said earlier the risk of diversion of a medicine in a countable pack (as opposed to e.g. a liquid in a stock bottle) is close to zero and would be rapidly identified even if only one single pack disappeared.

page 30 3.17: We fully recognise that depression can be episodic and agree with the clinical expert but our aim is always remission and conventional antidepressants have a significant relapse prevention effect ("The average rate of relapse on placebo was 41% compared with 18% on active treatment; Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Geddes et al, Lancet 2003;361:653-61). If treated well it does not need to be episodic.

page 21 3.19: This would best be phrased "Treatment-resistant depression can be an episodic condition if inadequately treated.

page 22 3.20: We would suggest a more accurate phrasing would be "treatment again if symptoms returned". I think you'll find this is true of any treatment and clinical area.

P31 3.30: That may be true but experience from abroad shows that the dissociative effects can be significant for an hour or so but that, after the first treatment, wear off with subsequent treatments and rarely need anything more than reassurance plus a low or managed stimulus environment. The ERG should not underestimate or overestimate this and it will require some training but not a huge amount. training is required in general to understand and managed TRD better.

page 31 3.30: We are concerned at the financial calculations in relation to the service impact. Secondary care mental health NHS Trusts are capable of safely managing CDs e.g. methylphenidate, methadone, buprenorphine, and opiates (all more abusable than esketamine) so the costs of managing a registry would be possible though we accept service pressures would need discussing and funding. Costings should also include savings from lower use of ECT which includes employing consultant anaesthetists and larger numbers of staff.

P32 3.31: use of an ECT suite: We are unsure what this means. Trusts will not necessarily need to use ECT suites but could find a quiet area with appropriate monitoring and more homely and calm than the probable clinical sterility of an ECT suite.

P33. 3.31: As highlighted in last response if 82% of MH Trusts have plans on how to implement esketamine this is a large proportion and gives an illustration of the need for an effective treatment for TRD. we agree this will be challenging however this is often the case in mental health services and feel not a reason to refuse a licensed treatment. Yes, it might take a bit longer but, as COVID-19 has shown, Trusts can implement changes extremely quickly when they want to. That only 18% did not have plans might now be out-of-date and is a low number. therefore an extended implementation would be needed for some STPs

P36 3.34: This is true. Methadone, methylphenidate and buprenorphine are examples.

P6 3.3: ECT personal comment: ECT is an option but is not a passive treatment. A colleague at Minds entire job at one time was supporting people pre-ECT and post-ECT with their fears, apprehension, distress, and memory loss. ECT also has a number of important contraindications and cautions:

ECT: Contraindications

Definition
Before discussing contraindications, it is important to first understand the physiologic effects of ECT. These include:
- Large increases in cerebral blood flow and intracranial pressure
- Initial parasympathetic discharge manifested by bradycardia, occasional asystole, premature atrial and ventricular contraction, hypotension and salivation
- Following parasympathetic reaction is a sympathetic discharge associated with tachycardia, hypertension, premature ventricular contractions, and rarely, ventricular tachycardia and ECG changes, including ST-segment depression and T-wave inversion, may also be seen.
- Glucose homeostasis is also affected. Hyperglycaemia seen in insulin dependent patients

Absolute contraindications:

- Known pheochromocytoma

Relative contraindications: The risk of the patient's psychiatric illness, side effects of antidepressant medications must be weighed against the risk of ECT and anaesthesia. These conditions include:

- Increased intracranial pressure, ok if there is not a mass effect
- Brain tumours, same recommendation as above
- Recent stroke- ECT has been performed successfully
- Cardiovascular conduction defects. Pacemaker is not a contraindication to ECT- AICD function can be deactivated and magnet should be available if needed
- High-risk pregnancy- OB consult and fetal monitoring is recommended
- Aortic and cerebral aneurysms
- Asthma/COPD- some suggest that you should discontinue theophylline because of its potential to cause status epileptics.

Recommendations:

- Delay ECT for patients with unstable angina, decompensated heart failure, or severe symptomatic valvular disease until these conditions are stabilized or optimized. Cardiology consultation may be of benefit
- For high-risk neurosurgical lesions including recent stroke and brain tumour, neurosurgical consultation is recommended
- Diabetic patients should hold oral hypoglycaemic, short acting insulin and halve their long acting dose with fasting
- Warfarin can be continued in high risk patients with INR <3.5
- In severe GERD antacids can be taken or intubation considered.

(Source https://www.openanesthesia.org/ect_contraindications/ accessed 10.9.20).

Are the recommendations sound and a suitable basis for guidance to the NHS?

We feel that this TA could read as if the chair and panel has decided to not recommend esketamine regardless of any evidence.

Concentrating on uncertainty is a technique regularly used in politics and business, and was used by e.g. the tobacco industry to try to minimise the effects of overwhelming data (but not 100.00% certain) that smoking caused lung cancer. This emphasising of "doubt" was effective and allowed the tobacco industry to keep up sales (and deaths) for many decades. A similar campaign is being waged on climate change because no one can be 100.00% sure it's human caused and, if it isn't, we can carry on as we are. We think you are asking too much for a novel and innovative therapy for a life-threatening condition in the early stages of introduction. It is easy to come up with uncertainties but we like the committee to attempt to see through these and see the overall message.

we would accept that esketamine could be 4th or 5th line in TRD (which is what will likely happen anyway) or that you give esketamine a limited approval with a further review in 2 years' time when more is known about the longer-term outcomes and more studies are published, and the

registries show up, but would ask it is not turned down entirely because of the inevitable uncertainties about the finer details of the studies. Studies in such heterogenous conditions do not always come up with blanket treatment effects as you might for a more homogenous illness.
Esketamine may be an on-going treatment for some, involving a day, a visit to a clinic (possibly many miles away, especially in the many rural areas), the need for an accompanying person or taxis, a treatment that is rather more than just popping a pill, and a significant routine. Not dissimilar to oncology treatments in fact and that can be for life-threatening conditions too.
We would thus appeal to the ERG to just take a step back and think "we have a condition here that is so horrible people commit suicide, we have a treatment that clearly helps a significant number (even if there are some uncertainties about the deeper details), and the alternatives are lacking or crude and with low patient acceptability. Can we really deny severely depressed people a treatment that works when other treatments do not?"

P24 3.22: Personal comment: A pharmacist in the UK who after 11 years TRD had decided to end it all as she could no longer bear the daily depressive symptoms but, after 3 treatments of esketamine (her last roll of the dice), phoned her husband to tell him where her secret stash of poisons was kept and told him to destroy them as she no longer wanted to end her life.

response prepared on behalf of CMHP by [redacted]

Name	[redacted]
Comments on the ACD:	

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
There is adequate real- world evidence in the USA/ Canada where this useful treatment has been available for almost 2 years. We have clinical evidence from some centres such as Oxford and Northampton in the UK. I have listened to many podcast by other experts highlighting its benefits in treatment resistant depression. At the moment with all of our combination drug treatments there is a significant "time lag". Only ECT can "jump start" recovery but relapse rates are high and some patients need maintenance ECT for a long time. I believe that this "rapid" acting antidepressant with "high remission" rates at 4 weeks is similar to ECT at 4 weeks of biweekly treatment. This is far higher than in STARD level 3 (14%) and or 4 (13%). Our patients and we as clinicians should not be deprived of this novel break through treatment option. I am writing as an advocate of my patients who are not very vocal. I am hoping that the patient groups and professional organisations will make similar points. SMC has already approved its use in Scotland within its marketing authorisation. Why should English TRD patients be deprived of this excellent treatment option as an alternative or an intermediate step before ECT.?? If cost is an issue acute and long term then at lease restrict its use but not deny it to our vulnerable patients as a

useful treatment option. Where has the patient choice gone? This is a major public health disorder with serious disability and personal consequences. In the health economic modelling you will need to take into account expensive and lengthy hospital stays for ECT and or other combination treatments. There must be some cost savings as this novel treatment can be administered on an out patient basis with home treatment support.

Are the recommendations sound and a suitable basis for guidance to the NHS?
Certainly not. In my personal opinion it should be available after 2 AD's have failed at adequate doses and adequate duration with good compliance. Depending on the severity of the illness. Very severely ill patients who might be considered for ECT must have this is an alternative. Less severely ill patients can go down the switch to Vortioxetine (within license) or augmentation treatments with Atypicals and or Lithium to start with. Esketamine could become a 4th line option before ECT.

Name	[redacted], The Royal College of Psychiatrists
Comments on the ACD:	

Has all of the relevant evidence been taken into account?
Evidence on the efficacy and safety of ketamine should be considered. The neuromodulatory effects of esketamine in the brain are likely to be identical or extremely similar to ketamine at equivalent doses.
In this regard, the highest level of impartial evidence is likely to be Cochrane Reviews. These show
that evidence in support of ketamine in depression is generally of poor quality, involving small samples, and with efficacy only shown over brief (clinically likely to be irrelevant) time periods. Also, risk of bias was often unclear, due to a lack of reporting. See:

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011612.pub2/full?highlightAbstract=ketamine%7Cketamin

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011611.pub2/full?highlightAbstract=ketamine%7Cketamin

Some academic institutions have patents and other intellectual property regarding ketamine and/or esketamine. This is more a case in the USA. Such conflicts of interest are often not mentioned when members of those institutions give presentations as to the apparent benefits of ketamine and/or esketamine.

Comment on Esketamine is likely to be used later in the treatment pathway because it has a higher treatment burden than other treatments
We agree that the position of esketamine in the treatment pathway is initially likely to be at least fourth or fifth line – i.e. after trials of augmentation. We also agree that for some patients this is because of the burden of treatment. Patients may not drive following esketamine treatment until they have had restful sleep. They can return home using public transport when they are fully recovered.

However, we consider that this later use also reflects its expense, novelty and association with a drug of abuse, more than the clinical data. Compared with the alternatives it is not obviously less safe. Therefore, particularly once costs come down, and particularly for patients are well supported, it is likely to be used earlier in the pathway. For some of those who are less well supported it may be more appropriate to provide better hospital or volunteer transport than to withhold the medication until later.

Comment on the response and remission evidence from TRANSFORM-2 should be considered with caution because of the short duration of the trial
We agree that 28 days does seem short as a primary end point for trials of antidepressants that are typically taken for many months. However, we do not think it is right to say that this has 'little bearing' on the treatment of depression. This is the internationally agreed time frame for licensing trials because of the ethical difficulty of leaving people on placebo for longer. Uniquely amongst programmes for a new antidepressant, the short term 28-day data in TRANSFORM are supplemented by the high quality data of the 1 year study SUSTAIN 2. Usually, it is lower quality post licensing studies that are used to clarify longer effects. It would not seem reasonable to withhold this drug from widespread use on the basis of a criticism that can be levelled at all other antidepressants which are in current use.
'The committee acknowledged that splitting the data into 2 groups could have inflated the differences between arms, particularly because the mean reduction in MADRS was near to the threshold for response in both arms at day 28. So, people could meet the criterion for symptom response in 1 arm but only have minimal differences in MADRS score in the other arm'.

We do not agree that splitting the data into 2 groups could have 'inflated the difference'.
The fact that the difference in remission, which is based on an absolute threshold level of the MADRS, between the two arms in TRANSFORM 2 (21.5%) is greater than the difference in response (17.3%), which is dependent on change relative to baseline level, effectively disproves the possibility of an inflated effect.

We would further make the point that response and remission are entirely conventional, pre-specified, measures. This new concept of a 'threshold in response' does not make sense when the difference in MADRS needed to meet criteria for response will vary for each participant depending on baseline. It is no more right to make decisions based on a NICE-generated post-hoc analysis which suggest that an effect size is 'near to threshold' than it would be to make decisions based on company-generated post hoc analyses which showed big effects.

Comment on the TRANSFORM-2 study is not powered to detect difference in effect between treatment arms so could show a false positive result
We do not understand why universally accepted standards for accepting a difference between two arms of a trial, are described as potentially a 'false positive'. It is of course possible that any result could be a 'false positive', but this is why we have accepted norms of statistical significance. The language here seems inappropriate. One would not accept a comment from a company which

asserted that their non-significant result was potentially a false negative because a study was underpowered!
The powering of the study is based on the number of patients to detect a difference assuming a specific degree of variance. It is possible that the difference was statistically significant despite the smaller than estimated effect size because, even though the difference was smaller, the degree of variance was lower.

Comment on withdrawal effects are difficult to distinguish from symptoms of depression
We agree that it is difficult to conclusively disprove that a new symptom arose because of stopping the drug rather than because relapse. However, the pattern of new symptoms provide important evidence as to which was happening and this does not appear to have been considered. In SUSTAIN 2 (Wajs et al 2020 Supplementary 5) the following effects were common (all >20% in the second week after cessation): insomnia, anxiety-nervousness, dysphoric mood-depression, fatigue – lethargy – lack of concentration, irritability, difficulty in concentration. These are all symptoms of major depressive disorder. By contrast the following symptoms were much less common: loss of appetite, nausea - vomiting, diarrhoea, poor coordination, sweating, tremulousness, dizziness-lightheadedness, headache, muscle stiffness, weakness, increased acuity sound smell touch, paraesthesias, depersonalisation-derealisation. With the exception of loss of appetite, these are not features of major depressive disorder. The dominant problem is therefore more likely to be relapse in depression rather than new symptoms occurring due to a change in physiology induced by the drug.

Increased feeling of hopelessness on withdrawal are an important problem, but are much more likely to be due to relapse in depression rather than being caused by the drug.
The short acting nature of the drug means that if it did induce some sort of change of physiology which caused withdrawal symptoms, then these effects would be expected to occur between each weekly dose (thereby undermining its beneficial effect), rather than after the end of a course. This was not observed.

For these reasons, we think the results of SUSTAIN 1 should be taken at face value.
The main implication of SUSTAIN 1 is that the drug needs to be taken continuously to prevent relapse. It undermines the company's assertion, made on the basis of much less direct evidence, that relapse will not occur if it is withdrawn later.

Comment on the differences in relapse rate in the SUSTAIN-1 trial data should be considered with caution
There seems to be a disparity between the conclusion – that the results of SUSTAIN 1 should be treated with caution – and the text which follows, all of which seems to point to reasons why the data of an outlier should not be excluded. The choice of language here seems inappropriate.

Comment on the evidence for esketamine is limited in its generalisability to the NHS

Severity

We agree that the trial data are limited in the degree of generalisability to populations that are more severe, but do not think this is a strong argument against adoption. Current practice is to use the same antidepressants in people with depression of all severities. The choice of antidepressants at different points on the treatment pathway is determined by side effect profile rather than by different antidepressants having different efficacy in different severities.

Comorbidities

The poor generalisability associated with the exclusion of patients with comorbidities is also relevant, both to safety and efficacy. Most psychiatric disorders, are associated with depression; and each will sustain and fuel the other. This is a contributory factor to high rates of prescribing of antidepressants in the population. We think the appropriate way to manage the risks of prescribing in patients with comorbid illness is through good phase 4 studies following adoption, rather than by withholding the drug from people with 'pure' resistant depression because it might be used in people with complicating comorbidities.

This can work in unexpected ways. For example, there are data from multiple studies suggesting that ketamine can be of benefit in reducing substance misuse. Clearly, however, there are also risks in people who are vulnerable to developing addiction, as reflected in the datasheet.

Comment on it is not appropriate to adjust the efficacy estimates of the placebo arm in the trials

Whilst we agree with the company's assessment of the influences on the placebo effect, we agree with the committee that the sort of post-hoc adjustment which the company applied was not appropriate.

Comment on safety must be considered when administering and monitoring esketamine

We agree with the committee that a registry is required. Further, we consider that such a registry should be interrogatable. Otherwise, those wishing to prescribe other rapidly acting antidepressants (eg IV ketamine) cannot be sure whether an individual is additionally taking esketamine nasal spray and is 'topping up'.

This would be a first step on the way to the use of systems such as Safescript (now mandatory in Australia) and Drug Prescribing Monitoring Programmes (as used in every state of the US) which require that, before prescribing, doctors intending to prescribe certain scheduled drugs must interrogate databases to ascertain existing and previous scheduled drug use.

We agree with the committee and the regulators that the signal is not, at present, strong enough to justify withholding the drug from the larger number who may benefit. An interrogatable registry will help in tracking the extent of suicidal behaviour associated with relapse or non-response to esketamine. This is a phase 4 task.

Comment on economic model
3.17 – 3.19 The company's economic model does not reflect the course of the disease

We agree that there are 'minimal long-term outcome data for people with treatment-resistant depression' to inform modelling. The higher cost of the drug makes this more important than it is for other cheaper oral antidepressants.

Comment on the effect of subsequent treatments is underestimated and the ERG's adjustment is more plausible
We agree with the committee that clinical practice would not be that 3 treatments would be attempted within 12 weeks as each successive treatment failed. Cycling between treatment takes much longer than this.

Comment on the cost of a course of esketamine treatment may be underestimated
The committee is concerned about variations in the dose and frequency of treatment. This data already exists. The company's data, as submitted to the FDA, shows that a higher proportion of those who remit but do not respond take maintenance esketamine at the shorter, weekly maintenance interval (69%) than those who remit (34%). In other words, those who respond less well take it more frequently.

Comment on A 1 to 2 ratio of nurses to patients is an appropriate resource cost during post-administration monitoring
We disagree slightly with the committee here. Based on the experience of the 5 UK centres which administer IV ketamine - for which the recovery time and requirements are likely to be similar if not slightly higher than for nasal esketamine - we consider that a 1 to 3 ratio more accurately reflects the need for healthcare staff supervision. Post treatment observation can be done by a healthcare assistant and, depending on the layout of the clinic needs only to be intermittent rather than continuous. It does not require a qualified nurse.

We agree with the clinical expert that the staffing need will change as clinics develop experience and efficiency of procedures. A typical ECT department would be able to start by treating esketamine patients at the end of their twice weekly ECT lists, thus avoiding employment of new staff until numbers justified a new bespoke clinical session. Fairly quickly, a single nurse and healthcare assistant can run a clinic with 3 concurrent patients each of which will be in clinic for about 2 hours in total.
In a clinic which has the beds/chairs to manage 3 simultaneous patients, two staff would be comfortably able to treat 6 patients in a session, including time for recording notes and, depending on its complexity, completing the registry. It is important to note that, as with directed observed administration of other CDs, a doctor does not need to be immediately present for the treatment.

Comment on significant investment will be needed to use esketamine in the NHS, but costs are difficult to quantify

Based on our experience, we think the only physical infrastructure likely to be required in an ECT suite is a Controlled Drug cabinet. In other settings it may also be necessary to purchase suitable comfortable chairs.

The processes for transporting drugs to the ECT exist already and are part of routine hospital transport systems so this does not incur new costs. The arrangements for disposal of used devices consist of putting a bespoke bin (like a large blue sharps bin) in the department which, when full, is transferred back to pharmacy for formal disposal of the remnants of the devices. This uses existing transport arrangements and again is low cost.

Training: The procedure is not complex, training materials are provided by the company and this could be accomplished within existing allocation of training time.

Comment on it will take time and resource use for esketamine to become part of clinical practice

We agree that esketamine is potentially disruptive to existing practice but observe that this may be a good thing. For example, patients with resistant depression commonly find that they become disillusioned with CMHT services because, however good the support, their condition does not change (by definition). When they have a treatment which abruptly helps, their care rapidly aligns with the service which provides it. In our experience, this commonly then results in the CMHT wishing to discharge the patient. The service providing esketamine then finds itself with a rapidly increasing caseload of patients who, if they relapse, are potentially at high risk. One way of managing this risk is to have shared care with the CMHT, but this duplicates effort and can seem pointless to the patient. A better solution may be to draw the resource into the new service from the old. This sort of disruption is to be welcomed – but, like all disruption, may initially be unpopular.

We agree that esketamine services should not be confined to ECT services and that community settings would be suitable. However, the infrastructure – a clinic with comfortable chairs, separated by curtains, which is suitable for administration and recovery - is common in many NHS settings.

We also agree that the reality of NHS processes is such that the lead times of 6-12 months for implementation quoted are realistic. However, this is driven by institutional barriers to introducing new technologies. Because the 'technology' is very simple, private clinics will be much quicker in set up.

In conclusion, we would not describe the costs of setting up a clinic as 'substantial'. The staff running costs could reasonably be estimated as a third of a session of a band 6 and a band 3 nurse – about £20 per patient per treatment.

Name	[redacted]
Comments on the ACD:	

General Comments
As a Consultant Psychiatrist in the NHS, frequently treating patients with treatment resistant depression (TRD), I am extremely disappointed with your decision to not recommend Esketamine with a serotonin selective reuptake inhibitor (SSRI) for this patient group.

The most comprehensive study of MDD treatment resistance was the National Institute of Mental Health-funded Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. (ref 1) . In this study, patients with MDD underwent a series of sequential antidepressant treatments in monotherapy or combination, or psychotherapy trials using evidence based antidepressant treatment strategies.

Acute remission rates decreased with each STAR*D level (level one 37%, level two 31%, level three 14% and level four 13%). Resistance to treatment becomes markedly increased at level 3 (after failure of two treatments), and predicts a poor prognosis with respect to future treatment efficacy and tolerance. Further more, patients with TRD experience relapse at a higher rate than do those with treatment-responsive MDD. Even when patients with TRD respond to treatment, the overall relapse rate while continuing treatment with the same antidepressant is high after 2 (65%; within 3.1 months) and 3 failed trials (71.1%; within 3.3 months).(Ref 1) .

Therefore, there is a substantial unmet need for effective treatments that can sustain antidepressant benefits for this population with)TRD . New treatments with good evidence in short-term and in maintenance trials , such as Esketamine with a selective reuptake inhibitor (SSRI), are very much needed in clinical practice today in order to bring hope to our patients and alleviate their suffering.

I very much hope this decision is changed.

Conflict of Interest/Disclosures: I receive only a salary from a full-time NHS post. I have no shares or positions in the pharmaceutical industry. I have received in the past consultancy fees from most of the pharmaceutical companies based in the UK.

Esketamine for treatment-resistant depression

ERG response to ACD

Produced by Kleijnen Systematic Reviews Ltd (KSR)

Authors
Robert Wolff, Deputy Director, KSR
Nigel Armstrong, Health Economist, KSR
Steve Ryder, Health Economist, KSR
Debra Fayter, Systematic Reviewer, KSR
Gill Worthy, Statistician, KSR
Caro Noake, Information Specialist, KSR
Jos Kleijnen, Director, KSR, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Robert Wolff, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 30/11/2020

Source of funding:
This report was commissioned by the NIHR HTA Programme as project number 12/78/96.

Declared competing interests of the authors
None.

Acknowledgements
Annette Chalker contributed to the data extraction and critique.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.

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

Contributions of authors
Robert Wolff acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Nigel Armstrong acted as health economic project lead, critiqued the company's economic evaluation, and contributed to the writing of the report. Steve Ryder acted as health economist on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence, and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report, and supervised the project.

Abbreviations

ACD	Appraisal Consultation Document
AiC	Academic in confidence
BSC	Best supportive care
CBT	Cognitive behavioural therapy
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CiC	Commercial in confidence
ECT	Electroconvulsive therapy
EMA	European Medicines Agency
EQ-5D-5L	European Quality of Life-5 Dimensions – 5 levels
ERG	Evidence Review Group
ESK	Esketamine
ESK-NS	Esketamine nasal spray
FDA	U.S. Food and Drug Administration
HCRU	Healthcare resource use
HDRS	Hamilton Depression Rating Scale
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
KSR	Kleijnen Systematic Reviews
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	Major depressive disorder
MDE	Major depressive episode
MHRA	Medicines and Healthcare products Regulatory Agency
MHT	Mental health trust
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NS	Nasal spray
OAD	Oral antidepressant
PAS	Patient access scheme
PHQ-9	Patient Health Questionnaire – 9 questions
QALY	Quality-adjusted life year
QIDS	Quick Inventory of Depressive Symptomatology
QoL	Quality of life
TRD	Treatment-resistant depression
UK	United Kingdom

Table of Contents

Table of Tables

1. Regulatory authorities assessed the ESK-NS clinical data (TRANSFORM-2 and SUSTAIN-1) and concluded they are robust and demonstrate the clinical value of ESK-NS

1.1 to 1.3 1) The TRANSFORM-2 results clearly indicate a statistically significant and clinically relevant treatment effect and outcome for patients with TRD; 2) TRANSFORM-2 was sufficiently powered and well-controlled, and not associated with a risk of a false positive finding; 3) Response and remission are established and appropriate outcomes in depression, and the four-week duration of TRANSFORM-2 is appropriate and is aligned with clinical trial design guidance from the CHMP

The company argued that the difference of LS mean change of -4.0 in MADRS observed in TRANSFORM-2 was meaningful and was unlikely to be a false positive finding, i.e. to have occurred by chance. The company also cited the EMA and TA367 to support its view that response and remission measured at 4 weeks are appropriate. On that basis they argued that four weeks follow-up is sufficient for decision making.

The ERG would concur that this is probably the case in the sense that the 95% CI did not overlap the point of no difference. However, the assertion that the result might have occurred by chance was made in the ACD in the context of the use of a difference of -6.5 in the power calculations, which suggested that a difference of -4 might not be of clinical significance and thus the finding might be a false positive if at least -6.5 was the minimum value to indicate a 'positive' result. As stated in the ERG report, the ERG would maintain that, given the uncertainty in difference in outcomes between ESK-NS+OAD and OAD over at least the duration of an episode of TRD, which is almost certain to last longer than four weeks, four weeks follow-up is insufficient for decision making. It might be the case that response or remission can be determined at 4 weeks, but that does not mean that patients should not continue to be followed up. To be most useful to decision making patients in trials should follow as closely as possible the care pathway that would be expected to be observed in clinical practice. Therefore, those patients who had not responded should have progressed to the next line of therapy and continued to have been followed up.

1.4 The random withdrawal design of SUSTAIN-1 is the commonly recommended approach for a long-term maintenance trial by health authorities, and additional regulatory analyses conducted concluded that unblinding did not impact the robustness of the trial results

The company state that the design of SUSTAIN-1 was recommended and accepted by 'many health authorities' for 'relapse prevention trials', citing the FDA and EMA. The ERG would agree that, in principle, if the decision was whether to discontinue ESK-NS + OAD or not in patients who had responded to or gone into remission on ESK-NS + OAD, this would be the appropriate design i.e. to randomise such patients to continue or discontinue ESK-NS + OAD. However, the decision problem includes patients who are ESK-NS inexperienced, as on entry to the TRANSFORM-2 trial. Unfortunately, by curtailing the TRANSFORM-2 trial at 4 weeks i.e. be transferring those who had responded or gone into remission into the SUSTAIN-1 trial, it was no longer possible to continue to follow up beyond 4 weeks the patients in both arms to which patients had been randomised. This meant that there was only an unbiased estimate of response or remission for ESK-NS + OAD vs. OAD up to 4 weeks. Beyond this there was no unbiased estimated of relapse or loss of response for ESK-NS + OAD vs. OAD in those patients in the index population of the decision problem i.e. those inexperienced to ESK-NS. By not following up both arms in the TRANSFORM-2 study, the estimates of relapse and loss of response had to come from two different sources, those for ESK-NS + OAD from SUSTAIN-1 and those for OAD from another source, the one chosen being STAR*D, with the inherent risk of bias

that is entailed. Also, as stated in the ERG report, relapse was only counted in those who had become 'stable remitters', defined as *"…a MADRS total score of ≤12 for the last two weeks of the optimisation phase plus for at least three of the last four weeks of the optimisation phase with one excursion of the MADRS total score >12 or one missing MADRS assessment permitted at Week 13 or 14 of the optimisation phase only"* (p.117). This means that some patients who would have been counted as remitters because of having a MADRS score ≤12 at 4 weeks post-randomisation in TRANSFORM-2 might not have been included in the analysis of relapse because they had not had a sustained MADRS score of ≤12. Such patients, given their lack of sustained MADRS score might have been more likely than those with a sustained score to relapse, which means that the rate of relapse on ESK-NS + OAD, which was used in the cost-effectiveness model, might have been underestimated.

1.5 Patients with suicidal ideation were not excluded from ESK-NS trials and patients with high risk of suicide were studied in a separate clinical development program.

In the ACD response, the company refers to its response to NICE's Draft Technical Report. In response to these comments, the ERG stated that *"patients with (previous) but not current (within 6 months) suicidal ideation / intent were included in the esketamine trials. Patients with suicide behaviour in the 12 months prior to the study were excluded. The company stated that "the TRD population studied is representative of a population with increased risk of suicidality". However, the exclusion of patients with "acute suicide risk" remains of concern to the ERG. The company have conducted a separate clinical programme in patients with a moderate to severe depressive episode of MDD who have current suicidal ideation with intent. Patients with TRD were not excluded from these trials. The results of these trials, when published in full, may be informative in relation to patients with TRD and at "acute suicide risk"."* This response still applies.

1.6 A robust risk management plan has been agreed with the MHRA

The ERG notes that *"Janssen is working closely with the MHRA to finalise the protocol for the registry and shares monthly reports of orders of ESK-NS, with the purpose of a continuous monitoring of the potential for abuse"*. Once finalised, this registry should be carefully assessed by all relevant stakeholders, including the clinical expert and the NHS commissioning expert who discussed the need of a registry.

1.7 Results from an additional long term safety study show no unexpected safety signals

In the response to the ACD, the company highlights *"interim unpublished data from the long terms safety study SUSTAIN-3 [showing] that there were no new safety concerns identified with continued intermittent ESK-NS dosing of up to 30 months (54 [4.7%] patients) as compared with the already determined safety profile in patients exposed to ESK-NS for up to one year"*.

The ERG discussed interim results of SUSTAIN-3 before (see section 4.2.9 of the ERG report), i.e. interim safety results from a cut-off of 31 December 2018 which included data from 1,140 patients treated for a mean of 13.7 months, and noted three deaths in SUSTAIN-3. As highlighted in the ERG report, *"this study, when reported in full, will give a fuller picture of any potential longer-term risks with ESK-NS including those related to withdrawing from treatment"*.

2. The outcomes predicted by the economic model are reflective of the outcomes that patients with TRD experience in the long term and the proportion of patients in MDE health state is appropriate, especially when a revised method for subsequent treatments is incorporated.

2.1 The health states used in the model are appropriate, established and based on previous NICE depression models

The company stated that the model structure was adequate. The ERG would not disagree that this might be the case in principle given that all models are only an approximation of the real world and notwithstanding the suitability of sources of parameter estimates. However, as stated in the ERG report, the ERG continue to assert that the effectiveness of subsequent therapy was probably underestimated. This is therefore liable to lead to a rate of movement to the MDE state with no possible remission that the committee believed was too high. The company also stated that they had addressed this concern by changes to the model in terms of the modelling of the effectiveness of subsequent treatments (see section 2.3, which also contains the critique of this method by the ERG).

2.2: Results from a targeted literature review of patients with TRD shows that long term outcomes of patients with TRD are poor

The company conducted a *"targeted literature review"*, summarised in Appendix I to show how poor the long-term outcomes of those with TRD are. The results of the studies in this review were shown in Table 3 of the company document. In this table there is also a comparison with the outputs from the company's revised model, including a revised method of subsequent treatments, see section 2.3. On this basis, the company conclude that the cost-effectiveness analysis should be considered conservative.

The ERG can confirm that the target literature review seems to show estimates of remission that are not dissimilar to those in STAR*D, although the ERG cannot locate the figures of 4.85% and 3.76%. The main limitation of all of the studies in the literature review is the lack of long term data in the context of TRD being an episodic condition, i.e. according to the low rates of remission it appears that patients would be left in the MDE health state for the remaining life expectancy. The ERG has not been able to conduct a review of the literature, but one study that the ERG have located provided an analysis of the treatment journey of patients with MDD, including those considered to have TRD (Wu 2019, doi: 10.1371/journal.pone.0220763). In this study routinely collected data on patients in the USA on the period at least 12 months before and after MDD diagnosis were analysed to estimate the duration of the episode and the time spent in remission between episodes, as well as number and type of treatments. The mean duration of the first TRD episode was estimated to be 571 days and that of the second was 482 days. The mean duration of remission for those who had TRD was 330 days. Of course, it cannot be assumed that everyone with a first episode of TRD will have a second and that the second episode will be TRD. Indeed, the number of 2nd TRD episodes was tiny in comparison to first, i.e. 93 vs. 3,317, the duration of a 2nd treated MDD episode was only 183 days and of remission for non-TRD was 407 days. However, even if every patient continued to have MDEs, half of which were TRD and half of which were not, for their remaining lifetime then this would imply a ratio of approximately (482+183)/(482+183+330+407) = 51% of lifetime spent in the MDE state. According to the model using the company revised method for subsequent treatments (see sections 2.3 and 5 for details), the ratio is 10.49/13.83 = 76%. Using the original ERG method that produced the ACD 2 preference values this ratio would be 52%. Using the method as part of the ERG scenario 1 that responds to the implausibility of the values for subsequent treatment relapse (see section 5), the value would be 48%.

2.3 Given the Committee's concerns with the proportion of people in the MDE state, a revised method for subsequent treatments is proposed which reduces the proportion of patients in the MDE state over time

The company stated that they adopted the revised ERG subsequent treatment method. This is in contrast to the original ERG method as described previously (see section 5 for details).

3. The use of the base case MDE utility is appropriate, and an alternative approach which addresses the Committee's concerns using amended criteria for MDE and a different utility value is provided for consideration

The company argue that the MDE utility is appropriate to use for those who relapse. However, this assertion by the company is largely a response to a problem identified by the committee regarding the MADRS threshold used to determine health states. This can be best illustrated by the discrepancy between the MADRS scores of patients who relapse and those of patients at baseline in the TRANSFORM-2 study from who the MDE utility is estimated. Specifically, to be eligible for TRANSFORM-2 patients had to have a MADRS score of at least 28 and the mean value was 37.1 (Table 4.8, ERG report). This threshold was also employed for the definition of the MDE health state in the model. However, relapse was defined as MADRS of at least 22 and [redacted] This means that some patients (with MADRS between 22 and 28) were counted as being relapsed would not have been defined as being in the MDE state according to either the eligibility criterion for the data source for the utility estimate for MDE or the definition of the MDE health state. This means that the rate of relapse on ESK-NS + OAD might be considered to have been overestimated in the model due to this discrepancy.

The company presented some additional evidence of utility estimates for TRD that were either similar or lower than that used for the MDE health state in the model. It also argued that utility of the MDE state is likely to decline over time, thus implying that its approach, which assumes no decline, is conservative with respect to the effect of relapse and thus the benefit of ESK-NS.

The ERG would concur that in principle that employing the MDE utility on relapse appears to be valid in that one would assume that those who relapse return to the MDE health state. The ERG would also agree that the value of the utility of the MDE health state, based on the evidence presented by the company albeit selective, is probably no higher than that used in the company base case. Although this does not seem to address the fundamental problem of the estimate of relapse based on the lower threshold of MADRS of 22 as opposed to 28, the discrepancy might have led to a higher estimate of relapse for ESK-NS + OAD. However, this is notwithstanding the high risk of bias in the relative estimates of relapse between ESK-NS + OAD and OAD given the different source for each, the former being from SUSTAIN-1 and the latter from STAR*D, as referred to in section 1.4.

3.1 Amending the criteria for MDE health state allows consideration of an alternative utility from a QoL study conducted in UK patients with TRD

Despite its defence of the approach to estimating the utility of patients on relapse, the company have presented a scenario analysis where the utility of MDE was raised from 0.417 to 0.430, which was the mean EQ-5D-5L value from the UK TRD Quality of Life study of patients with a PHQ-9 score ≥13.8, which the company state is equivalent to a MADRS of at least 20, see section 5.1 for results. Only a small amount of additional information on the TRD Quality of Life study was reported in Appendix D, but the full report of the study had already been presented in response to the technical report.

The ERG consider that it is reasonable to consider different values of utility for the MDE state. However, the ERG would question the value of employing a lower estimate of utility for the MDE state given the company's additional evidence, albeit selective, that the base case value might be too high. Indeed, it is worth noting that the PHQ-9 score threshold employed to indicate the population of interest, described as TRD, in the UK TRD Quality of Life study was 17.8, stated to be equivalent to a MADRS score of 28, which is consistent with the definition of MDE in the model. This led to an estimate of 0.354 for EQ-5D, i.e. lower than the base case value. Also, the value for remission estimated based on a threshold of PHQ-9 of 9.4, stated to be equivalent to a MADRS of 12, was also lower than that estimated for the base case, which was 0.866, based on the same MADRS threshold. Indeed, it was also lower than the value for response of 0.764, which was based on a MADRS of greater than 12, patients only having to have had at least a 50% reduction in MADRS. The ERG consider that, if it is valid to consider employing the utility value from the UK TRD Quality of Life study, it is probably also valid to employ the value for remission. However, it is not clear to the ERG that amending the utilities in the model will address the concerns of the committee regarding MADRS thresholds.

4. It is appropriate to assume different healthcare resource use costs per health state, which could lead to different medical costs between treatment arms. We propose a sensitivity analysis with reduced cost differences among health states to address the Committee's concerns and include additional costs that may be associated with commissioning ESK-NS

4.1 It is not appropriate to use SUSTAIN-1 to inform HCRU per treatment arm and evidence provided shows differential costs per health state is appropriate

The company defended their approach to estimating non-drug costs according to health state as opposed to the committee's preference for assuming equal *"healthcare resource use costs across treatment arms"*. It also argued that SUSTAIN-1 is not an appropriate source of resources by which costs can be calculated, largely on the basis of it not having been designed to collect these data and thus the short follow-up time once relapsed i.e. in what might be regarded as the MDE health state.

4.2 Sensitivity analysis with reduced cost differentials between health states

The company have performed a scenario analysis that use the lower 95% CI limit instead of the mean value of cost from the company submission data source, see section 5.1 below for results.

The ERG would agree with the company that non-drug costs should be estimated in a state transition model as a function of health state. The ERG would also agree that SUSTAIN-1 is probably not a good source of resource use data. However, the higher hospitalisation rate with ESK-NS + OAD as opposed to OAD, albeit based on very few and limited data, does raise concerns regarding the lower non-drug costs due to less time spent in the MDE state with ESK-NS + OAD estimated in the cost-effectiveness model. The method of using the 95% CI lower limit does seem to be one reasonable alternative, although it does not overcome any underlying problems with the method of estimating the costs used by the company. The ERG would also argue that notwithstanding problems with the precise method of estimating health state costs, the problem of underestimating the effectiveness of subsequent therapies, as mentioned in issue 2, could be partly responsible for any overestimation of the non-drug cost difference between ESK-NS + OAD and OAD. However, what does seem particularly perplexing is that, although the company cite two previous economic analyses performed for NICE, CG90 and TA367, to support their claim of a cost differential between the MDE and remission health states, both the differential and absolute values reported in these studies are much lower than in the company model. Both of these studies make use of a costing study by Byford 2011, which was large (n=88 935 patients), longitudinal, conducted in the UK and designed to estimate the effect of remission vs. non-remission. Although the population was not limited to those with TRD, the annual healthcare cost in the severe depression subgroup was £749 in remission vs. £1,037 in non-remission, which translate into 28-day costs of £57 and £79, respectively. These can then be inflated using the NHS cost inflation index (NHSCII) to give £63 and £87. This contrasts with 28-day costs of £164 and £980 for remission and MDE respectively in the company model. One might argue that TRD is more costly than severe, but it does seem questionable that the difference would be so large, i.e. by a factor of about 11 for MDE vs. non-remission in the inflated figure from Byford 2011. On this basis the ERG have conducted further scenario analyses using the inflated Byford 2011 values (See section 7).

4.3 Estimates of the costs of commissioning are incorporated into the model and have very minimal impact on the cost effectiveness

The company reported the results of a survey of 16 mental health trusts (MHTs) of the commissioning cost of implementing ESK-NS services, most fully described in Appendix E of the ACD response. The results seem to show that the 16 respondents expect little if any additional increase in cost associated

with providing facilities to support delivery of ESK-NS. However, the total costs for all 16 MHTs for controlled drug cabinets was found to be £2100. The company calculated that, if was extrapolated to cover all 69 MHTs and split by the total expected number of patients who will be treated with ESK-NS in the first 5 years, it would result in a cost of £1.62 per patient. This cost was then incorporated in their scenario analyses (see section 5 below).

5. Revised post ACD- 2 scenarios for consideration (Full population)

The company have provided two additional scenarios for the index TRD population, as shown in Table 6 in the company ACD response. One of these, referred to as the 'Janssen revised ACD scenario', included the ERG revised method of estimation of subsequent treatments (including BSC) first introduced in the ERG critique of the ACD response by the company. This is based on the assumption that, in order to calculate the values of remission and response for subsequent treatment (including BSC), i.e. TRD lines 2 to 5, the values for remission and response at TRD line 1 are those from STAR*D, i.e. 13.7% and 16.8%, as opposed to those from TRANSFORM-2, i.e. 26.6% and 18.4%, as in the original ERG method presented in the ERG report. The company presented the results for these scenarios alongside what they described as the *"NICE ACD 2 preferred scenario"* in Table 6. However, the ERG could not match the figures of [redacted] (* indicates excluding carer disutility) to the results of any previously submitted documents. Values were reported in the ACD that might be regarded as the NICE ACD 2 preferred scenario, but they were not the same: *"Using the committee's preferred assumptions, the ERG's ICER was in the range of £64,554 to £72,158 per QALY gained, including no carer disutility and the ERG's carer disutility, respectively"* (p.34). According to Table 6, there were only three differences in assumptions between the 'NICE ACD 2 preferred scenario' and the 'Janssen revised ACD scenario'. In the latter, differential medical costs (referred to above as 'non-drug related') per treatment arm base on health state costs instead of equal medical costs per treatment arm, the effect of subsequent therapy was estimated using the original ERG approach and commissioning costs were included, the result of which was an ICER range of [redacted] The other additional scenario presented in Table 6 was named 'Janssen revised ACD 2 scenario: sensitivity analysis', where two additional changes were made to the 'Janssen revised ACD scenario', i.e. utility value for MDE from TRD QoL study used (see section 3.1 above) and lower bound of 95% CI for health state costs used (see section 4.2 above). The ICERs for this scenario are: [redacted]

The ERG could not reproduce the results of the company in the version of the model, which was used to produce the ICERs reported in the ACD for the second ACM, according to committee preferences and which the ERG can set to produce the revised company base case ICER of [redacted] as well as the ICERs based on the ACD 2 preferences. This model will be referred to as the 'ACD 2 model'. The ERG have checked the implementation of subsequent treatment effectiveness in the company model and found the factor for calculating reduction in risk of response and remission between treatment lines after TRD line 1 (point of receipt of either ESK-NS + OAD or OAD) is as stated by the ERG in both the original and revised method i.e. 16.3/16.8 for response and 13.0/13.7 for remission. There is, however, a discrepancy between the method as implemented by the ERG and as implemented by the company in the value assumed for TRD line 1 used to estimate subsequent lines. The ERG assumed that the value was that for TRD line 1 as expressed in STAR*D, i.e. as step 3, which is 16.8% for response and 13.7% for remission. However, the company have taken a weighted average of TRD line 1 and line 2, as expressed as step 3 and step 4 respectively in STAR*D, which is of 16.8 and 16.3 for response and 13.7 and 13.0 for remission. The weights are 62.39% for TRD line 1 and 37.61% for line 2. The ERG also noted that the company had also applied this weighting to calculate the risks of relapse and loss of response for subsequent treatment. This weighting is according to the percentage of patients in SUSTAIN 1 who were at lines 3 and 4, corresponding to TRD line 1 and 2, respectively. The ERG would therefore concede that this weighting approach is probably more valid and therefore any subsequent ERG scenario analyses have been implemented in this way. However, the ERG's proposed method of adjustment by using a factor calculated from the ratio of the risk at step 4/risk at step 3 from STAR*D has created values for the risk of loss of response and

especially the risk of relapse that appear to be implausibly high (See Table 1 and as shown in Appendix C of the ACD response).

Table 1: Subsequent treatments: risk of loss of response and relapse

	Original ERG method	Revised ERG method, as implemented by company	Original CS
Loss of response			
TRD Line 2	23.0%	23.1%	22.8%
TRD Line 3	23.7%	23.7%	22.8%
TRD Line 4	24.3%	24.5%	22.8%
BSC/ Non-Specific Treatment Mix	25.0%	25.2%	10.4%
Relapse			
TRD Line 2	17.4%	17.1%	12.8%
TRD Line 3	32.7%	32.3%	12.8%
TRD Line 4	61.6%	61.0%	12.8%
BSC/ Non-Specific Treatment Mix	116.0%	99.0%	4.2%

Indeed, the value for BSC (TRD line 5) for relapse is impossible in that it lies above 1 and the original company values for BSC were both lower than for earlier lines of subsequent treatment. This makes no sense in the context of the motivation for the ERG adjustment i.e. because subsequent treatment effectiveness was perceived to be too low in the original CS. It is because of these implausible values that the ERG have conducted an additional scenario analysis that caps the risks of loss of response and relapse at 22.8%% and 12.8% respectively, which are still higher than in the original CS (see section 7).

The ERG also note that, according to Appendix C, the company replaced the OAD relapse value of 9.2%, which was stated to have been estimated from SUSTAIN-1, with 9.0%, which was stated to have been estimated from TRANSFORM-2, with no obvious motivation. The ERG have found that, even when applying all of the changes listed in Appendix C and Appendix F of the ACD response, none of the results shown in Table 6 could be reproduced in the ACD 2 model. However, the results for the Janssen revised ACD scenario can be reproduced in the ACD 2 model by replacing the value for OAD loss or response of [redacted] with the value in the latest company models of [redacted]. Doing this would then lead to a value of [redacted], for the revised company base case [redacted]

5.1 Revised scenario at later line positioning: non-response to at least 3+ prior OAD.

The company presented results of analysis of TRANSFORM-2 and TRANSFORM-3 of the subgroup of only those patients who did not respond to at least 3 prior OAD treatments (3+ subgroup), which were presented for change from baseline in MADRS in Table 7 and for remission and response in Table 8 of the ACD response.

[redacted]

[redacted]

The company also performed a new cost-effectiveness analysis for the 3+ subgroup where the changes in parameters were stated to have been listed and results shown in Table 16 in Appendix G. The results of this analysis, referred to as 'Janssen revised ACD scenario at later line positioning (non-response to at least 3 prior OADs)', were an ICER range of [redacted].

The ERG were able to reproduce the results of the 'Janssen revised ACD scenario at later line positioning (non-response to at least 3 prior OADs)' in the ACD 2 model, although there was a very small difference, i.e [redacted], which seemed to be associated with a discrepancy in life years only beyond the 9th decimal place. This required changing the parameter inputs as shown in Table 2, as well as utility and dosing values.

Table 2: Comparison between parameter values in the 3+ line model vs. original model

Parameter type	Treatment /line of therapy	3+ line as required to produce the reported ICER range.	All patients/original model
Remission	ESK-NS + OAD	[redacted]	0.461
	OAD+PBO-NS	[redacted]	0.266
Response	ESK-NS + OAD	[redacted]	0.155
	OAD+PBO-NS	[redacted]	0.184
Relapse	ESK-NS + OAD	[redacted]	0.056
	OAD+PBO-NS	[redacted]	0.092
Loss of response	ESK-NS + OAD	[redacted]	0.042
	OAD+PBO-NS	[redacted]	0.224
Remission	Subsequent treatment TRD line 2	12.1%	12.8%
	Subsequent treatment TRD line 3	11.5%	12.1%
	Subsequent treatment TRD line 4	10.9%	11.5%
	BSC	10.3%	10.9%
Response	Subsequent treatment TRD line 2	3.5%	3.4%
	Subsequent treatment TRD line 3	3.7%	3.5%
	Subsequent treatment TRD line 4	3.8%	3.7%
	BSC	3.9%	3.8%
Relapse	Subsequent treatment TRD line 2	23.7%	23.1%
	Subsequent treatment TRD line 3	24.4%	23.7%
	Subsequent treatment TRD line 4	25.2%	24.4%
	BSC	25.9%	25.2%
Loss of response	Subsequent treatment TRD line 2	32.1%	17.0%
	Subsequent treatment TRD line 3	60.6%	32.1%
	Subsequent treatment TRD line 4	99.0%	60.6%
	BSC	99.0%	99.0%

All of the values in Table 2 could be located or calculated from those reported in the company response to ACD 2, except for Relapse OAD+PBO-NS in the company 3+ line model, which were not as reported in Table 21, i.e. 0.127942040 and 0.228134772 respectively. Utility and dosing values for the 3+ subgroup for TRANSFORM 2 and 3 were not reported at all, but only located in the company 3+ line model.

Note that the values for subsequent treatment and BSC used for 3+ line are precisely the same as those in the original model except that they apply to a relatively earlier line of therapy e.g. 12.1% is applied to TRD line 3 in the original model, but TRD line2 in the 3+ line model. This can be explained by considering that the lines of therapy are all in effect one line later in the 3+ model given that line 1, where ESK-NS + OAD is administered, would be given to patients who had failed on one more treatment than in the original 'all patients' model. The ERG considers that this method of shifting lines of therapy does make some sense, but it also compounds the problem of the implausibility of the values of loss or response with a value of 99% for both line 4 and 5 (BSC).

In Appendix H of the ACD response the company also presented scenarios based on the Janssen revised ACD scenario, but using either the lower or upper bounds of the 95% CIs for dosing estimated in TRANSFORM-2, TRANSFORM-3, and SUSTAIN-1. The effect was to either decrease or increase the ICER range from [redacted] to [redacted] or [redacted], respectively.

6. Other points and Errors/ factual inaccuracies

Table 11 of the ACD response is reproduced as Table 3 below. A column 'ERG comment' has been added.

Table 3: Other points and factual inaccuracies

Location in ACD and statement	Rationale	ERG comment
Section 3.12 *'the company also did not use data from SUSTAIN-1 for relapse rate in the oral antidepressant with placebo arm in the economic model to avoid any withdrawal effect.'*	STAR*D was used to inform the relapse rate in the OAD arm to avoid the potential bias that could occur when using the SUSTAIN-1 OAD + PBO-NS data due to its ESK-NS withdrawal study design. However, as an alternative scenario, data from SUSTAIN-1 has been used to provide alternative estimates or relapse and loss of response for OAD (Section B.3.4.4.8 of company submission). This decreases the ICER compared to the use of the STAR*D data.	The ERG agree with the company that data from SUSTAIN-1 were used in a scenario analysis, although, as the ACD states, such data might be contaminated by a withdrawal effect from patients having been in stable remission on ESK-NS immediately before randomisation to OAD + PBO-NS.
Section 3.5 *'The ERG added that the network meta-analysis only used adjusted effects for the oral antidepressant with placebo arm of esketamine.'*	The network meta-analysis using the unadjusted effects were provided to NICE in July 2019 in Section D.1.3.4 of the appendices and the Company response to Question C2 of the ERG Clarification Questions.	As stated in the ERG report (see section 5.2.4), the ERG acknowledged that NMA results with TRANSFORM-2 effectiveness unadjusted for placebo effect was provided in response to clarification.
Section 3.6 *'An expert from the NICE guideline on depression noted that psychological therapies were not included as comparators or with combination treatments in the company's submission but were included in the NICE appraisal scope.'*	This is an error, as psychological therapies were not included as a comparator in the NICE appraisal scope.	Psychological therapies are not included in the NICE final scope.
Section 3.8: *'The committee also noted that the score used for relapse was not equivalent to the MADRS score for moderate to severe depression, which affected*	A score of 20-34 on the MADRS scale is regarded as moderate severity. The threshold used for relapse, a MADRS score of 22 or higher, is	Not a factual inaccuracy.

Location in ACD and statement	Rationale	ERG comment
the health state utility values and transitions in the economic model.'	consistent with presence of MDE symptoms of moderate severity.	
Section 3.15: *'The committee concluded that it had not seen evidence that the additional clinical contact involved in the placebo arm improved clinical outcomes.'*	This is not aligned to the evidence provided, which shows that the improvement in clinical outcomes from clinical contact does not rely on CBT to improve clinical outcomes. This was demonstrated in the Posternak study.	Not a factual inaccuracy. The ACD summarised the considerations by the committee and as well as the conclusion reached.
Section 3.17: *'The model output suggests that within 1 year, 78% of people with treatment-resistant depression in current clinical practice do not have symptom response to any treatments long-term. So, they then occupy the MDE state for the remainder of the time horizon.'*	The statement regarding that patients occupying the MDE health state for the remainder of the time horizon is incorrect, as patients who are in the BSC treatment phase have an ongoing likelihood of achieving response (3.9% chance per 4 weeks) or remission (10.5% per 4 weeks) and hence a proportion of patients will continuously move out of the MDE state.	The ERG agree that this is misleading: it is true that there remains a non-zero probability of remission and response. However, the rates if relapse and loss of response are high enough such that after about a year the rate of increase of the proportion in the cohort on BSC in the MDE state is approximately the same as the rate of increase of the proportion in remission or relapse. Therefore, on BSC patients spend about 17 years or 85% of the 20-year time horizon in the MDE state.
Section 3.21: *'The transitions between response and remission states were also sourced from STAR*D for both arms, although this assumption was not fully explored by the company.'*	This is a factual inaccuracy. Data from SUSTAIN-1 were used to inform the rate of transition from response to remission, as noted in Section B.3.2.9.2.1 of the company submission.	The ERG agree that this is a factual inaccuracy.
Section 3.21: '*The relapse and loss of response rates for the oral antidepressant arm were sourced from the STAR*D trial. The STAR*D trial used different relapse criteria. Also, it was unclear if the population from STAR*D is generalisable to the NHS.'*	Relapse is defined as a return of the MDE following the achievement of remission but before fulfilling the criteria for recovery from the current episode. The definitions of relapse used in STAR*D and SUSTAIN-1 are appropriate to capture this worsening of depressive symptoms. In STAR*D, relapse was declared when the QIDS-SR16 score collected by the interactive	This is not a factual inaccuracy.

Location in ACD and statement	Rationale	ERG comment
	voice response system during the follow up phase was ≥11 (corresponding to an HRSD17 ≥14). This definition correlates to moderate to severe depression, which is similar to the criteria in the SUSTAIN-1 trial.	
Section 3.23: *'The committee noted that the mean EQ-5D-5L health score index was consistently higher than 0.8 for all participants at the end of maintenance for SUSTAIN-1. In participants who were randomised to withdraw from esketamine, 45% of people whose depression was in stable remission and 58% of people whose depression was in stable response relapsed. The committee considered that this would not correspond to the relatively high EQ-5D-5L health score index above 0.8 if this represented a true transition to the MDE health state.'*	The SUSTAIN-1 utility data for stable responders and stable remitters do not capture patients who transitioned to the MDE health state. SUSTAIN-1 only included stable remitters and stable responders. If patients relapsed or lost response, then they no longer contributed to these data sets. This can be seen as the sample size in the SUSTAIN-1 stable remitters and stable responders reduces over time. Patients who are in stable remission or stable response can be expected to have consistently higher utility scores, and the SUSTAIN-1 study data are consistent with this.	The ERG agrees with the company that those utility values above 0.8 do seem to apply to only those patients who continued to be in a state that was defined as either stable remitter or stable responder (see CS, Table 31).
Section 3.27 '*The committee would like to see the proportion of people having each dose, how often people have esketamine (weekly or every 2 weeks), reasons for the dosing choices.'*	Please see Appendix H for further scenarios exploring the impact of different dosing scenarios. The reasons for the dosing choices were previously provided to the ERG (see response to ERG clarification questions, Section A.9).	The proportion of people having each dose and whether weekly of two-weekly was not reported by the company. Instead the mean number of visits per week and devices per session were reported in Appendix H. The reasoning for the complex dosing regimen during the trial remained opaque to the ERG except that, according to the company, it was intended *"…to emulate real-world clinical practice"* (Response to request for clarification).
Section 3.28 *'The committee considered that CBT and ECT were excluded from the trials and should not be included in the medical costs.'*	CBT and ECT should not be excluded from the costs of the health states. Whilst the retrospective chart review showed that CBT and ECT do not comprise a large proportion of the costs of	The ERG agrees with the company that for the purpose of modelling the care pathway that it might be appropriate to include the cost of any treatment that TRD patients might experience over a 20-year time horizon.

Location in ACD and statement	Rationale	ERG comment
	patients with TRD, it is not appropriate to exclude these costs, which are still incurred in the NHS.	
Section 3.30 *'costs associated with creating and managing a registry to avoid misuse and abuse of esketamine.'*	Costs to the NHS will be minimal as the company has agreed with the MHRA to cover the costs of the data collection. Nurse time for the administration of the registry will be captured during supervision of the patient.	See section 1.6
Section 3.31 *'They said a reasonable time to implement esketamine in a community setting would be 12 months, and 6 months in a secondary hospital clinic setting.'*	Feedback from multiple mental health trusts indicates 180 days is not required. Feedback from NHS at a Trust level has clearly said that significant infrastructure investments are not required. 82% of the sites said that they will repurpose existing premises for the adoption of ESK-NS into the NHS.	Not a factual inaccuracy. ACD summarising advice received by an NHS commissioning expert.

7. Additional ERG analyses

Given the implausibility of the values for loss of response and relapse in the revised ERG method of subsequent therapy effectiveness, the ERG has conducted a scenario analysis starting with the committee's preferred assumptions, as reported in the ACD, which led to an ICER in the range of £64,554 to £72,158 per QALY gained, including no carer disutility and the ERG's carer disutility, respectively (See Table 4). This is first modified to produce the NICE ACD 2 preferred scenario given the new PAS. Following that the subsequent therapy effectiveness is incorporated using the company implementation of the ERG revised method plus the change in values for OAD relapse and loss of response to produce the Janssen revised ACD scenario. ERG scenario 1, which attempts to adjust for the implausibility of this method is then shown. ERG scenario 2 adapts ERG scenario 1 by using the inflated costs from Byford 2011 (See 4.1 above).

Table 4: ERG scenarios for whole population (based on the model that produced the values reported in the ACD with two changes^)

Key differential parameters	NICE ACD 2 preferred scenario	NICE ACD 2 preferred scenario (with new PAS)	Janssen revised ACD scenario (with new PAS)	ERG scenario 1: loss of response +relapse risks capped (with new PAS)	ERG scenario 2: Byford 2011 costs + Janssen revised ACD scenario
MDE utility	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)
Medical (HCRU) costs	Equal medical costs per treatment arm	Equal medical costs per treatment arm	Differential medical costs per treatment arm based on health state costs	Differential medical costs per treatment arm based on health state costs	Differential medical costs per treatment arm based on health state costs
Subsequent treatment approach	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$
Best Supportive care approach	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction, based on TRANSFORM-2 OAD arm efficacy	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$
Additional costs of commissioning	Not included	Not included	Included (see Section 4.3)	Included (see Section 4.3)	Included (see Section 4.3)
ICER	£64,554 - £72,158*	[redacted]	[redacted]	[redacted]	[redacted]

^using 9.0% instead of 9.2% for OAD relapse; [redacted] instead of [redacted] for OAD loss of response.
$using revised ERG method, as implemented by the company (see Section 5.)
*excluding carer disutility

As can be seen in Table 4, with the new PAS, the ICER might be above the £30,000 threshold depending on the assumptions made regarding subsequent treatment effectiveness and non-drug costs.

A similar set of scenario analyses was performed for the 3+ line subgroup and shown in Table 5. ERG scenario 3 shows the effect of introducing the method of capping the loss of response and relapse risks to those at TRD lines 2 to 4 as in the original CS, which is to increase the ICER, but not enough to be higher than the £20,000 or £30,000 threshold. Only if equal medical costs are assumed would the ICER exceed £20,000: still below £30,000 or above £30,000 if subsequent treatment effectiveness is estimated using the company implementation of the ERG method or the capped method respectively.

Table 5: ERG scenarios for line 3+ population (based on the model that produced the values reported in the ACD)

Key differential parameters	**Janssen revised ACD scenario at later line positioning (non-response to at least 3 prior OADs)**	**ERG scenario 3: loss of response +relapse risks capped (with new PAS)**	**ERG scenario 4: equal medical costs + Janssen revised ACD scenario at later line positioning (non-response to at least 3 prior OADs)**	**ERG scenario 5: Byford 2011 medical costs + loss of response +relapse risks capped (with new PAS)**
MDE utility	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)	TRANSFORM-2 baseline utility (MADRS total score of ≥28)
Medical (HCRU) costs	Differential medical costs per treatment arm based on health state costs	Differential medical costs per treatment arm based on health state costs	Equal medical costs per treatment arm	Differential medical costs per treatment arm based on health state costs
Subsequent treatment approach	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$
Best Supportive care approach	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$	STAR*D Step 3-4 reduction based on STAR*D efficacy$
Additional costs of commissioning	Not included	Not included	Not included	Not included
ICER	[redacted]	[redacted]	[redacted]	[redacted]

*excluding carer disutility
$using revised ERG method, as implemented by the company (see Section 5.)

ERG critique of company ACD 2 response: clarification of errors and factual inaccuracies

19th January 2021

Janssen wish to clarify the following questions from the ERG on replication of scenarios and statements, which are found below:

Statement from ERG	Clarification	ERG comment
P.7 The company argued that the difference of LS mean change of -4.0 in MADRS observed in TRANSFORM-2 was meaningful and was unlikely to be a false positive finding, i.e. to have occurred by chance.	A minor point, but Janssen wish to clarify that the value of -4.0 difference in MADRS is the difference between the LS means in the ESK-NS+OAD and the OAD+PBO-NS treatment arms in TRANSFORM-2, and not the LS mean change in MADRS from baseline within a treatment arm. The mean change in MADRS total score from baseline to the end of induction –21.4 for ESK-NS + OAD versus –17.0 in OAD + PBO-NS arm (p=0.010).	Not a factual inaccuracy. The ERG would like to thank the company for this clarification.
P.9 The ERG can confirm that the target literature review seems to show estimates of remission that are not dissimilar to those in STAR*D, although the ERG cannot locate the figures of 4.85% and 3.76%.	The figures of 4.85% and 3.76% are the probability of sustained benefit at Level 3 and 4 of STAR*D respectively. The figures are reported in Table 1 of Sackheim (2016): • Sackeim, H. (2016). Acute Continuation and Maintenance Treatment of Major Depressive Episodes With Transcranial Magnetic Stimulation. Brain Stimulation, 9(3), 313-319. doi: 10.1016/j.brs.2016.03.006	Not a factual inaccuracy. The ERG would like to thank the company for this clarification.
P.14 However, the ERG could not match the figures of [redacted] (* indicates excluding carer disutility) to the results of any previously submitted documents.	This is the ACD2 Committee preferred ICER, but with the new PAS (list price ICER of £64,554-£72,158, net price ICER of [redacted].	The ERG still cannot reproduce these figures of [redacted]) when applying the new PAS to £64,554-£72,158.
P15. The ERG have conducted additional scenario analysis that caps the risks of loss of response and relapse at 22.8%% and 12.8% respectively	Minor point: the capped values which are used in the cap for relapse and loss of relapse are not identical to the ones used in previous models (rounding to 1 decimal place). The actual value for relapse: 12.7942040149209% The actual value for loss of response: 22.813477196106%	Not a factual inaccuracy given that the values chosen should not be regarded as precise estimates of the risks.

P.15 The ERG also note that, according to Appendix C, the company replaced the OAD relapse value of 9.2%, which was stated to have been estimated from SUSTAIN-1, with 9.0%, which was stated to have been estimated from TRANSFORM-2, with no obvious motivation.	This was done because the starting cohort in the model is based on the TRANSFORM-2 population rather than the SUSTAIN-1 population. The company ACD 2 response (Appendix F) notes that the source to inform the proportion of patients who have failed 3 prior treatments included in the model. This has been amended from SUSTAIN-1 to TRANSFORM-2 in all scenarios in the ACD 2 response. This means that the original OAD relapse risk from remission to MDE (TRD line 1) has changed from 9.2% to 9.0% in the model.	Not a factual inaccuracy. It remains unclear to the ERG why the change from SUSTAIN-1 to TRANSFORM-2 was made.
P23, Table 5 ERG scenario 5: Byford 2011 medical costs + loss of response +relapse risks capped (with new PAS)	ERG Scenario 5: We are unclear on if either 1) the assumption of equal medical costs, or 2) differential medical costs using Byford 2011 have been used in this scenario. In ERG scenario 5, the ICERs of [redacted] cannot be replicated, as we are not clear what costs are used in this scenario. Minor point: ERG scenario 4 should be labelled as scenario 5.	This has now been corrected.

Esketamine nasal spray (ESK-NS) for treatment resistant depression (TRD) (ID1414): Company response to ERG critique of ACD2 response: 19th January 2021

We would like to thank NICE for the opportunity to respond to the additional scenarios and new evidence the ERG have presented. Once the two key points below (incorporation of ERG cap and differential medical costs) are considered further, the plausible ICERs and cost effectiveness estimate of ESK-NS within the optimised 3+ failure TRD population (patients who have failed at least 3 prior oral antidepressant (OAD) therapies) are within NICE's threshold for decision making.

Key Point 1) The ERG's new cap on the relapse rates of subsequent treatments has the overall impact of decreasing the proportion of patients in MDE over time. The Wu et al (2019) study is cited by the ERG to support this, however this study should be considered with caution because of generalisability to the UK and methodological issues. As such, the scenarios using the subsequent treatment relapse cap included in the ERG analyses (ERG Scenario 1,3,4) should be considered as the upper bound cost effectiveness estimate of ESK-NS.

The Wu et al study used as the rationale for the implementation of the cap by the ERG should be considered cautiously, as it is a retrospective analysis of US insurance claims, and therefore has low generalisability to the UK and has significant other issues. For example, using prescription data instead of clinical outcome data to inform relapse rates is likely to underestimate the duration of MDE. Additionally, selection bias during the follow up period means that patients with a prolonged and incomplete MDD episode during the follow-up were excluded. This would lead to an underestimated time in an MDE. The Wu et al estimates of MDE duration over a lifetime are much lower than the evidence from the targeted literature review on observational and randomised clinical trials (see Company ACD 2 response) that shows the long term outcomes of patients with TRD in the UK are very poor. Specifically, the best available evidence shows the relapse rate increases per each subsequent treatment line (Rush et al, 2006), which would result in a high proportion of patients experiencing MDE over time. We understand that at this stage of the appraisal there is limited opportunity to consider other sources, but we suggest to the committee that the ICERs with the implementation of the ERG cap should be considered the upper bound estimate of the cost effectiveness of ESK-NS, given the other sources identified in the targeted literature review to inform long term outcomes (Dunner et al, 2006, Sackheim, 2016, Fonagy et al, 2015, Aaronson et al, 2017).

Key Point 2) The Company and the ERG agree that differential costs per treatment arm based on health state costs are appropriate. We also note the Byford et al (2011) study used in one of the scenarios has been conducted in a primary care population and has low generalisability to a TRD population. The resource use and costs in the base case, from the cost study in a UK TRD population, is the more appropriate source than Byford et al, especially considering the optimised TRD 3+ OAD failure population.

There is agreement between the Company and ERG that differential costs per treatment arm, based on differential costs per health state, are appropriate (ERG critique, section 4.2). The rationale for this was summarised in the previous ACD 2 response (ACD 2 response, section 4.0). We acknowledge the ERG's rationale for considering the Byford et al study, as this has been previously used in the NICE appraisal for vortioxetine (TA367) and NICE clinical guideline models. A serious limitation of the Byford et al study, however, is that it has been studied in a broader and less severe MDD population, and is therefore not within the scope of the decision problem. Differences in the resource use and costs between the MDD and TRD population are well known (Mrazek et al, 2014, Jaffe et al, 2018, Johnston et al, 2019). Whilst the ERG do briefly caveat the difference in population between the TRD study and the Byford study, the significance of the use of a non-TRD population in this STA should be emphasised. The TRD cost study is more appropriate to inform the health state costs, given the population studied within the study directly matches the population for the decision problem (Denee et al, 2020).

Summary: Accepting these two key points, 1) taking the ERG cap as the upper bound estimate of cost effectiveness, and 2) excluding the equal medical cost assumption from the Byford study and ERG scenarios (Table 4 of the ERG critique), then the scenarios in the optimised 3+ OAD failure TRD population are below the NICE lower bound threshold of £20,000 per QALY and considered cost effective (Company ICER of [redacted]* and ERG Scenario 3, ICER of [redacted]*, including PAS).

We suggest that the inclusion of differential health state costs and using the UK TRD cost study is more appropriate, given the consideration of the ERG cap (which reduces the proportion in MDE over time, see Appendix A) and the optimised 3+ failure TRD population. All of these scenarios also include the additional costs of commissioning. Additional real world data collection in a UK population is planned (ESK-NS TRD post launch cohort study), which will increase the certainty of the cost effectiveness of ESK-NS. For this to occur, however, use of ESK-NS within the UK

setting is required. There are also further ongoing studies, the TRD3008 long term safety study and a TRD3013 phase 3B study comparing ESK-NS with OAD augmentation therapy.

We continue to believe that ESK-NS is a clinically and cost effective option for the NHS in the whole TRD population, but have explored the potential for an optimised recommendation in 3 or more OADs based on the ACD 2 and NICE engagement to provide further certainty on clinical effectiveness using the available data. Together with the increase to the PAS in October, it is hoped the committee are able to have confidence that the clinical and cost effectiveness case of ESK-NS has been demonstrated in an optimised population of patients in a specialist mental health setting with 3 or more OAD failures. Planned monitoring for regulatory authorities and data collection should give the committee confidence in recommending ESK-NS, while allowing clinicians and patients access to the first new treatment option and mode of action in depression for over 30 years.

*excluding carer disutility

References

Aaronson, S., Sears, P., Ruvuna, F., Bunker, M., Conway, C., & Dougherty, D. et al. (2017). A 5-Year Observational Study of Patients With Treatment-Resistant Depression Treated With Vagus Nerve Stimulation or Treatment as Usual: Comparison of Response, Remission, and Suicidality. American Journal Of Psychiatry, 174(7), 640-648. doi: 10.1176/appi.ajp.2017.16010034

Byford, S., Barrett, B., Despiégel, N., & Wade, A. (2011). Impact of Treatment Success on Health Service Use and Cost in Depression. Pharmacoeconomics, 29(2), 157-170. doi: 10.2165/11537360-000000000-00000

Denee, T., Ming, T., Waller, J., Bailey, T., Rajkovic-Hooley, O., & Middleton-Dalby, C. et al. (2020). A retrospective chart review study to quantify the monthly medical resource use and costs of treating patients with treatment resistant depression in the United Kingdom. Current Medical Research And Opinion, 1-1. doi: 10.1080/03007995.2020.1857580

Dunner, D., Rush, A., Russell, J., Burke, M., Woodard, S., Wingard, P., & Allen, J. (2006). Prospective, Long-Term, Multicenter Study of the Naturalistic Outcomes of Patients With Treatment-Resistant Depression. The Journal Of Clinical Psychiatry, 67(05), 688-695. doi: 10.4088/jcp.v67n0501

Fonagy, P., Rost, F., Carlyle, J., McPherson, S., Thomas, R., & Pasco Fearon, R. et al. (2015). Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock Adult Depression Study (TADS). World Psychiatry, 14(3), 312-321. doi: 10.1002/wps.20267

Jaffe DH, Rive B, et al. The Burden of Treatment Resistant Depression in Europe from the Patient Perspective. Presented at ISPOR, 10-14 November, 2018, Barcelona, Spain.

Johnston KM, Powell LC, et al. The burden of treatment-resistant depression: A systematic review of the economic and quality of life literature. J Affect Disord. 2019;242:195-210.

Mrazek DA, Hornberger JC, Altar CA, Degtiar I. (2014) A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. Psychiatr Serv. 65(8): 977-987.

Rush, A., Trivedi, M., Wisniewski, S., Nierenberg, A., Stewart, J., & Warden, D. et al. (2006). Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. American Journal Of Psychiatry, 163(11), 1905-1917. doi: 10.1176/ajp.2006.163.11.1905

Sackeim, H. (2016). Acute Continuation and Maintenance Treatment of Major Depressive Episodes With Transcranial Magnetic Stimulation. Brain Stimulation, 9(3), 313-319. doi: 10.1016/j.brs.2016.03.006

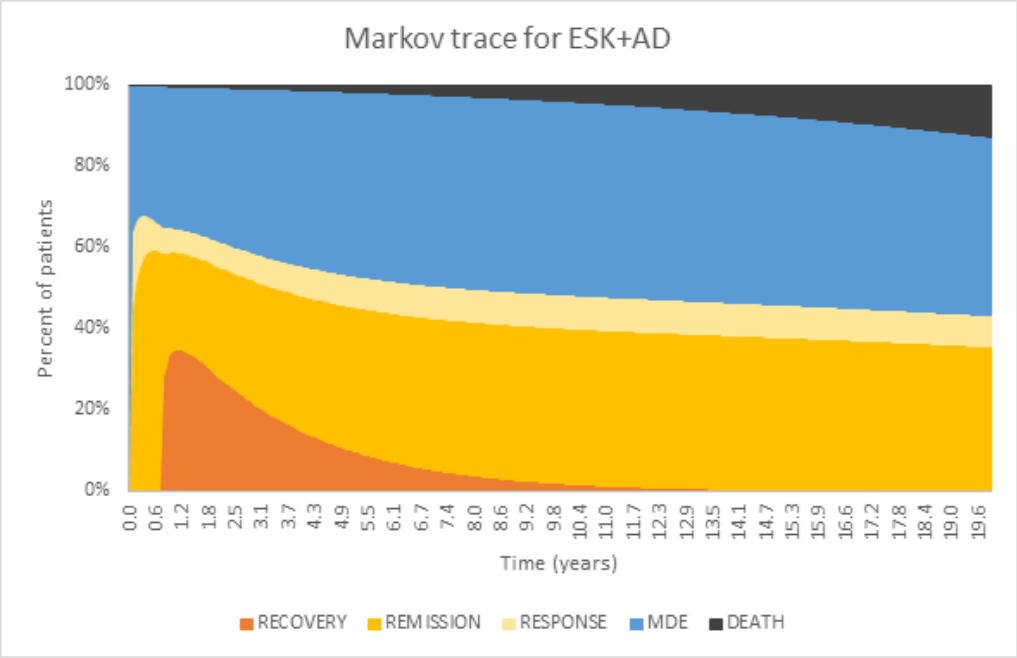
Wu B, Cai Q, Sheehan JJ, Benson C, Connolly N, Alphs L (2019) An episode level evaluation of the treatment journey of patients with major depressive disorder and treatment-resistant depression. PLoS ONE 14(8): e0220763. https:// doi.org/10.1371/journal.pone.0220763

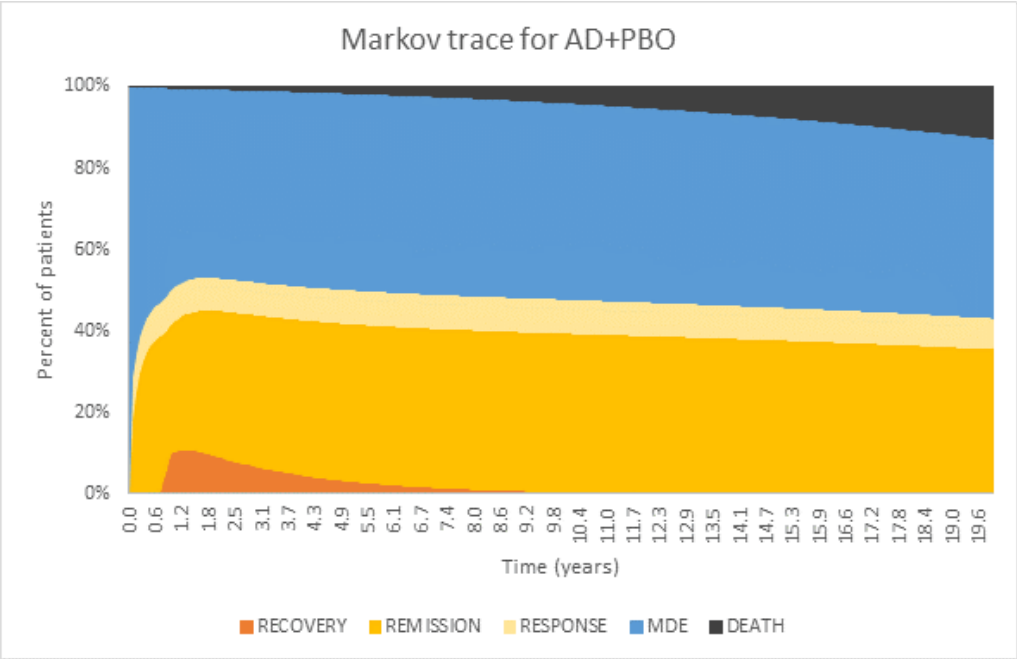
Appendix A: Markov Trace when ERG cap is implemented

The scenario using the ERG subsequent treatment relapse cap should be considered the upper bound of the cost effectiveness estimate of ESK-NS because the response and remission rates in the literature are lower than in the Markov trace.

When the ERG cap is implemented, the Markov trace shows that the proportion of patients in MDE over time is significantly reduced to 48%, which means there is a higher proportion of patients in remission and response over time compared to the literature.

The results from the targeted literature review (Appendix I of Company ACD 2 response) show the long term outcomes of patients with TRD are poor, with only low rates (<10-25%) of response (including remission), and also low rates (≤12%) of sustained remission (including recovery) are achieved in current clinical practice for up to 3.5 years. Therefore, the cost effectiveness estimate for ESK-NS should be considered conservative since the number of patients ending up in the MDE state over time is likely to be lower in the model than in clinical practice.





NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Esketamine for Treatment-resistant depression (ID1414)

Submission addendum

14th February 2022 (updated AIC/CIC May 2022)

File name	Version	Contains confidential information	Date
ID1414 esketamine for TRD Submission Addendum May Update AIC_CIC V2	V2	Yes	09/05/2022

Esketamine for treatment resistant depression [ID1414]

Janssen submission addendum – 14th February 2022

Executive Summary

Janssen thanks NICE and welcomes the opportunity to submit additional information to the Committee, and to comment on [redacted] esketamine for treatment resistant depression (TRD) [ID1414]. Janssen also wanted to thank NICE for the opportunity to explore alternative access routes for esketamine nasal spray (ESK-NS). [redacted]

[redacted] Janssen are returning to the NICE appraisal committee with several new elements for consideration based on discussion with NHSE&I and NICE. This consists of a revised commercial offer through a patient access scheme (PAS) [redacted], in addition to an updated value proposition, and an amended and optimised positioning of ESK-NS in the treatment pathway (after augmentation treatment). This is supplemented by providing the committee with new evidence that addresses some of the key uncertainties, which has become available since the 3rd appraisal committee meeting (ACM).

Seeking a NICE recommendation [redacted] allows for the most feasible and practical approach to generating additional evidence to address some of the committee's concerns. Janssen have previously stated the possibility of including a UK cohort within a pan-European post-access real world evidence (RWE) study or, ECHO clinical study, if the committee are able to recommend ESK-NS in one of the populations presented for consideration. We consider that ESK-NS is a cost-effective treatment for TRD in all TRD patients but take note of the concerns that the committee have raised with applying this to the English healthcare system; we also note many of these concerns are irreducible uncertainties without post-reimbursement data collection on the real-world use of ESK-NS. If the committee are able to recommend ESK-NS in one of the proposed optimised populations, it would allow for the collection of RWE, which would hopefully expand use of ESK-NS in future years by addressing the committee's uncertainty, while providing an additional option for TRD patients who have a high unmet need and for which clinicians have been without any new significant treatment options for the last 30 years.

Providing new treatment options in mental health is especially important given that mental health has been subject to significant under investment compared to physical health for some time, and the lack of parity of esteem between mental health and physical health is persistent and significant [1]. This situation has deteriorated since the outbreak of COVID-19 with several reports noting the decline in mental health and wellbeing of the world's population and in the UK. A historical lack of investment and research into mental health and depression means that inherent difficulties and irreducible uncertainties remain in the evidence base. Clinical trials in depression, and mental health generally, are notoriously associated with high placebo rates, plus patient and response

heterogeneity which makes it challenging to ascertain the true relative treatment effect of the active drug over placebo, resulting in challenges in conducting trials in mental health, which are acknowledged by regulators like the EMA. [2] [3] [4]The difficulties in the evidence base have also been highlighted in producing evidence-based recommendations for the new NICE depression guideline CG90, which were described as having 'potentially serious limitations.' [5] Irreducible uncertainty in the long term outcomes, which were noted by the Committee, means that an 'Only in Research' recommendation for ESK-NS would not resolve all uncertainties. Hence, given exploring a MAA route is not currently feasible, we hope the Committee are able to make a recommendation in baseline commissioning.

Over and above the challenges in generating robust evidence in mental health, there is a broader equality and equity consideration regarding access to innovative treatments in mental health. That is, the challenge of generating robust evidence means that there is a risk that new innovative treatments will remain underfunded, and patients suffering from mental health conditions will remain under-served, compared to, for example, cancer patients. This is most acutely highlighted by the fact that there have been no new treatments for depression with a novel mechanism of action for decades, which underlines the need to find solutions to bring new innovations to underserved disease areas, liked depression. Janssen remains committed to finding an access route for ESK-NS to provide another treatment options for people with TRD. In addition to the development program being the most comprehensive ever conducted in TRD, we are committed to ongoing research in depression, TRD, and for ESK-NS. This has led to supporting this appraisal through several bespoke studies and Janssen's plan to keep generating further evidence in this field, through studies like the ECHO study. Despite this continued investment in research, unfortunately it is not possible to fully address all uncertainties raised by the committee, for example the long-term outcomes of TRD, and we therefore ask the committee to be mindful and take into consideration the challenges of conducting mental health research, as well as the many open basic scientific questions regarding depression, especially for people TRD. We therefore suggest there is a need for acknowledgement that some uncertainty is currently irreducible, and pragmatism is required to ensure TRD patients with the highest unmet need are given an innovative treatment option.

Finally, we would also like to highlight the significant unmet need of all patients with TRD, but specifically for the two optimised populations we have proposed for NICE appraisal committee consideration. As noted previously, depression and TRD imposes considerable health and economic burden on the people with the condition and their families (including dependents and carers), in addition to the health service and wider society. People with TRD compared to MDD have increased mortality, reduced quality of life in the ranges of metastatic cancer or acquired blindness and higher cost of illness to themselves, and to the NHS and society through 50% lower labour participation and 20% increase in work activity impairment. [6] [7] [8] [9] [10] The disease burden is magnified at later lines of treatment, especially the burden on the NHS with higher use of resource utilisation and specifically expensive long term stays in hospital. [11] There is a need for new modes of action that can break the cycle of depression, especially for people who do not respond to current treatments given that remission and response rates decline with every treatment line. [12] Unfortunately, many people with TRD end up cycling through numerous ineffective antidepressants. In addition, many of the interventions that are used in later lines of therapy, like augmentation strategies which have very limited clinical evidence and significant adverse events, do not have a marketing authorisation

for use in depression and are therefore used off-license. The lack of evidence for these treatments is highlighted by the fact that in the network meta-analysis (NMA) for this appraisal, it was not possible to develop robust evidence synthesis and similarly, the NICE depression guideline update was not able to find evidence of these treatments in the appropriate position in the pathway either. [5] The widespread use of these off-license treatments in clinical practice highlights the unmet medical need for licensed treatments in TRD.

The revised optimised population is in people who have failed at least 3 oral antidepressants (OADs) and after augmentation. The updated positioning of ESK-NS meets the significant unmet need for people who have failed several previous treatment options, where there is a higher burden of illness for the those suffering with TRD, and their carers, as well as the NHS and wider society. [13] This position is also aligned to stakeholder, and clinical feedback received during the NICE consultation process, which highlights that later line patients may receive the most benefit from ESK-NS [14]. The revised optimised population represents approximately 10% of the eligible original TRD population (see appendix D). Furthermore, the 3 or more (3+) OAD failures and after augmentation population addresses many of the previously noted implementation and equity concerns discussed by the committee, given the considerably smaller population of patients who are largely managed in specialist secondary care mental health settings, where existing infrastructure is available to provide ESK-NS. This population is presented throughout this addendum and in a base case of the economic model, in addition to this, an optimised population, previously presented in the 3rd ACM, in TRD patients who have failed 3+ OADs is presented.

Overall, ESK-NS is the first antidepressant with a different mechanism of action in 30 years, which also received a breakthrough designation from the US Food and Drug Administration (FDA) because it is the first licensed fast-acting antidepressant indicated specifically for patients with TRD. Whilst the most used OADs primarily target the monoaminergic pathway, ESK-NS has a novel mechanism of action, targeting the glutamate pathway. It is thought to exert its action by N-methyl-D-aspartate (NMDA) receptor blockade, that is hypothesised to alter the underlying pathophysiological process of depression and have a neuroplastic effect. This unique mechanism of action results in a rapid onset of action (within 24 hours), and in combination with newly initiated OADs, provides greater response and remission rates for the short-term, and lower relapse rates for the long-term in comparison with currently available OADs used in the TRD population. [15] [16]

Summary of new evidence provided

Despite the innovation and clinical profile of ESK-NS which can help to meet the challenges associated with depression and TRD, we do understand the concerns raised by the Committee. We have therefore provided further information which addresses many of the concerns highlighted in the 3rd ACM. Furthermore, we have revised the base case economic model in line with the Committee's comments from ACM 3, as well as provided a new population (now 2 subgroups), value proposition and PAS for the Committee to consider. To summarise the new evidence which has been used to address the Committee's concerns and to present the new population and value proposition for consideration, please see the below summary of our response in the following sections:

1. Submitted new evidence to support ESK-NS in a new optimised population who have failed 3 or more OADs and after augmentation is aligned to clinical expert input on patients who have the highest unmet need and are most likely to benefit from ESK-NS.
2. New evidence from the long term ESK-NS safety studies shows that ESK-NS has a manageable safety profile, while emerging RWE demonstrates a consistent efficacy profile to the ESK-NS randomised controlled trials (RCTs).
3. A new analysis of UK data demonstrates the high resource utilisation and high healthcare cost of treating people with TRD, supporting the original TRD cost study rather than the Byford et al study, which is not appropriate given the setting of care and revised population.
4. The ERG treatment cap in the model and subsequent treatment efficacy is overestimating long-term outcomes for people with TRD based on the literature.
5. A new scenario is included in the base case for the Committee's consideration: overall, the new evidence and updated model shows that ESK-NS is a cost-effective option in both the 3+ OAD failures and the 3+ OAD failures and after augmentation position in the MDD pathway. The incremental cost effectiveness ratios (ICERs) with the new value proposition are below £10,000 per QALY threshold.

Table 1 Cost effectiveness results summary

Population	Total costs	Total QALYs	Incremental costs (ESK-NS vs comparator)	Incremental outcomes (ESK-NS vs comparator)	ICER
Sub population 1: at least 3 prior OAD failures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sub population 2: At least 3 prior OAD failures plus after augmentation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

In Appendix I, we have provided information regarding other updates to the company base case analysis based on available new evidence and the committee's preferred assumptions as outlined in the 3rd ACM. We note that many of committee's preferred assumptions were conservative based on the evidence presented.

Conclusion

Overall, in this addendum we wanted to reiterate and highlight the high unmet need for people with TRD and the impact this condition has on patients, families, and carers, alike; an impact that has become more pronounced since the COVID-19 pandemic. There is lack of equality and equity regarding access to innovative treatments in mental health compared to physical health. ESK-NS provides many benefits and meets an unmet need in helping people out of their depressive episode while supporting people with TRD, the NHS and wider society. We hope that the Committee can recommend ESK-NS in at least one of the populations presented, to provide a new option for patients and clinicians where other options have failed. If recommended, this will allow for a

manageable implementation of ESK-NS in the NHS and the possibility to collect post-reimbursement data, opening the possibility for a wider recommendation upon a NICE re-review of the evidence in the future.

1. Submitted new evidence to support ESK-NS in a new optimised population who have failed 3 or more OADs and after augmentation addresses a high unmet need and is supported by clinical opinion and available evidence (which demonstrates that ESK-NS would be a clinically and cost-effective option)

Janssen response:

It was clear from the clinical expert at the 3rd ACM and from some of the consultees during the 2nd ACD consultation that one of the positions suggested for ESK-NS was use after augmentation strategies had been tried. Following further discussions with NHSE&I and clinical experts in the field of depression, Janssen decided to present a new optimised population for reimbursement for people who had 3+ OAD failures and after augmentation, which is aligned to comments received and where most importantly there is the highest unmet need for people with TRD due to limited alternative available treatment options.

In summary, we would like to submit two populations for consideration by the committee at the 4th ACM:

- Base case 1: ESK-NS after the failure of 3 or more oral antidepressants (as presented at the 3rd ACM) – "3+ OAD failures"
- Base case 2: ESK-NS after the failure of 3 or more oral antidepressants and after 1 previous augmentation with an atypical antipsychotic or lithium – "3+ OAD failures and after augmentation"

The new optimised subgroup of people with 3+ OAD failures and augmentation will also help to manage previously raised implementation concerns, since eligible patients are a much smaller cohort of TRD patients and will be those who already require specialist secondary care mental health treatment with augmentation strategies, as outlined below:

1.1 The use of ESK-NS in the 3+ OAD failures and after augmentation subgroup fulfils a significant unmet need and is an appropriate position in the pathway for the committee to consider the clinical and cost effectiveness of ESK-NS

As people with depression move through the treatment pathway and cycle through different therapies, the available treatment options that offer therapeutic value become limited to the point where clinicians struggle to treat patients appropriately. Treatment options diminish to where the unmet need for this group of patients is significant. To further explore this group of patients who require treatment after augmentation, Janssen conducted an advisory board, where the clinical experts noted that there is a lack of understanding and knowledge around TRD within policy making and the medical community, who do not treat depression as a serious illness (Appendix B). The experts reported that patients with increasingly severe TRD are often forgotten and neglected, due to a loss of hope by the patient themselves and their clinician. This can affect patients' quality of life, relationships, and employment status. The clinical experts reported that many clinicians stop trying to find alternative treatment options, with some even re-diagnosing patients with other conditions after the failure of multiple treatments (Appendix B). Clinicians also resort to unlicensed or unconventional treatments at this stage of the pathway, given the lack of treatment options. [9] The impact on the patient and the economic effects of this suboptimal treatment are significant with many people ending in a situation of deprivation. (Appendix B) Additionally, the clinical experts agreed that these people are characterised as people who are generally silent about their treatment

and will not actively seek new treatment unless their carer or family member brings them to an appointment (Appendix B). Clinical experts, patient groups and previous NICE consultations have highlighted this population and consider the best use of ESK-NS to be in later lines of the treatment pathway. [14]

Individuals at these later lines have been shown to have a greater burden of illness than MDD patients. A recent UK study using the *Discover* dataset from Northwest London (NWL), captured healthcare resource utilisation between non TRD-MDD patients, TRD patients and those TRD patients at later lines of therapy. The study used a retrospective database analysis of the DISCOVER dataset, which is a longitudinal dataset covering over 2.5 million people who live and are registered with a GP in NWL, capturing all patients aged 18 and over with a diagnosis code for MDD in primary or secondary care, and a product code for an AD prescription during the study [redacted] The dataset is unique in that it links data from primary care, secondary care, community, mental health, and social care. The study included adults with (non-TRD) MDD and adults with TRD. [redacted] Table 2 shows [redacted] for patients [redacted] This is reflective of the depression [redacted] Overall, this results in an increasing burden to the NHS [redacted]

Table 2: Resource use by disease status [redacted]

Healthcare resource utilisation	MDD patients Mean (SD)	TRD patients (≥2 lines of therapy) Mean (SD)	TRD patients (≥3 lines of therapy) Mean (SD)	TRD patients (≥4 lines of therapy) Mean (SD)
Average number of OAD prescriptions	[redacted]	[redacted]	[redacted]	[redacted]
A&E Attendances	[redacted]	[redacted]	[redacted]	[redacted]
Non-elective admission	[redacted]	[redacted]	[redacted]	[redacted]
Elective admissions	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient first appointments	[redacted]	[redacted]	[redacted]	[redacted]
Primary care appointments	[redacted]	[redacted]	[redacted]	[redacted]
Mental Health hospital appointments	[redacted]	[redacted]	[redacted]	[redacted]
Referrals to secondary care	[redacted]	[redacted]	[redacted]	[redacted]

From the same study, Table 3 shows that the burden of illness for individuals at later lines of TRD [redacted]

Table 3: Burden of illness by disease status [redacted]

Burden of illness	MDD patients Mean (SD)	TRD patients (≥2 lines of therapy) Mean (SD)	TRD patients (≥3 lines of therapy) Mean (SD)	TRD patients (≥4 lines of therapy) Mean (SD)
Depression duration (years)	[redacted]	[redacted]	[redacted]	[redacted]
Average number of comorbidities of specified list	[redacted]	[redacted]	[redacted]	[redacted]
Suicidal ideation and attempt combined	[redacted]	[redacted]	[redacted]	[redacted]

Previous studies have also shown an increase in mortality and unemployment, as well as decrease in work activity and quality of life as the number of lines of failed treatment increases. [17] In addition to this, a higher impact on carers for these patients, whose carer burden is already considerable. [18]. Overall, the burden of illness for TRD is substantial, and increases at later lines of TRD and highlights the value of a treatment with a new mechanism of action being available to relieve the depressive episode at this stage of the treatment pathway.

1.2 To inform this new revised positioning, evidence of the efficacy of ESK-NS+OAD in the 3+ OAD failures subgroup has been adjusted to a population that have failed 3+ OADs and augmentation subgroup using available evidence and an approach previously accepted by NICE

In this section, we have presented additional evidence from SUSTAIN-2, which was an open-label single-arm safety study of ESK-NS where the direct entry patients had similar flexible dosing as the TRANSFORM-2 trial patients. In the base case analysis of patients who have failed 3+ OADs and augmentation:

- We have used the proportional ESK-NS treatment effect seen in SUSTAIN-2 between lines of patients that had failed 3+ OADs and had 3+ OAD failures and augmentation.
- This proportional ESK-NS treatment effect has then been applied to the ESK-NS data from 3+ OAD failures subgroup from the TRANSFORM-2 and 3 trials to generate an estimate of ESK-NS effectiveness for the 3+ OAD failures and after augmentation population.
- For comparator efficacy, we have very conservatively maintained the 3+ OAD efficacy from the TRANSFORM-2 and TRANSFORM-3 trials (i.e., not adjusted comparator efficacy)
- See Table 4 below for the inputs used, see Appendix C for further details

Table 4: Revised proportional reduction inputs from SUSTAIN-2 data for 3+ OAD failures and augmentation model base case inputs

Population/Treatment effect	Intervention	Remission	Response
3+ OAD failures model inputs (TRANSFORM studies)	OAD+PBO-NS	[redacted]	[redacted]
	ESK+OAD	[redacted]	[redacted]
Relative treatment effect between SUSTAIN-2 3+ prior OAD and 3+ prior OAD and augmentation subgroups, applied to ESK-NS+OAD efficacy only in case base		[redacted]	[redacted]

3+ OAD failures and after augmentation: Inputs used in the base case model	OAD	[redacted]	[redacted]
	ESK-NS +OAD	[redacted]	[redacted]

There is limited evidence available for comparator treatments for the 3+ OAD failures and augmentation population, as has also been identified by the committee in [redacted]. We therefore consider that the best available evidence of the short-term effect for a comparator remains the OAD+PBO control arm evidence from the TRANSFORM 2 and 3 clinical trials. Conservatively, we have not adjusted the absolute treatment effect for the likely less effective, later line of treatment for OAD+PBO as discussed in Section 1.3. This leads to a highly conservative estimate of the ESK-NS+OAD relative treatment effect vs OAD+PBO given the RCT evidence from TRANSFORM-2 (see below).

Janssen, as noted above, have adjusted the treatment effect downwards for ESK-NS+OAD despite evidence showing an improvement in relative treatment effect for ESK-NS+OAD vs OAD+PBO from the TRANSFROM-2 subgroups. Conversely, for OAD+PBO, we have taken the highly conservative assumption not to adjust treatment effect downwards, despite the significant drop in efficacy seen in OADs from the TRANSFORM-2 trial, where remission dropped from [redacted] between 2+ OAD failures and 3+ OAD failures lines, a relative drop of [redacted]. The equivalent decrease for people responding between these lines was [redacted] a relative drop of [redacted] We consider that the conservative nature of this assumption is also supported by the clinical experts' view at the 3rd ACM regarding 'the difference in [relative treatment effect between] subgroups were plausible and would be expected, because a newly started oral antidepressant would likely be less effective in the 3+ OAD failures treatments subgroup than in the 2+ treatments subgroup.' By inference the same could be assumed by moving from one line later from the 3+ OAD failures subgroup to the 3+ OAD failures and after augmentation subgroup for an OAD comparator and given the similar mechanisms of action for these agents. This inference was supported by the clinical experts that we consulted in the advisory board informing this submission addendum. (Appendix B)

In addition to evidence from TRANSFORM-2 and clinical expert opinion, evidence from STAR*D also suggests a declining treatment effect for OADs used at a later line. In STAR*D, the effect between Step 3 and Step 4 with remission was 13.7% and 13.0%, and 16.8% and 16.3% for response, respectively. Although this was a small reduction in efficacy versus that observed in TRANSFORM-2, this could be hypothesised to be a result of the introduction of combination and augmentation treatments being tried at this stage of the STAR*D trial. [12] Fifty-nine patients had augmentation and combination treatments (out of the 123 patients) at Step 4 of the trial, which may suggest a stabilising of treatment effect with augmentation. To reflect this, we have conservatively maintained the efficacy of the OAD 3+ inputs from the TRANSFORM studies, rather than reducing downwards further. We also know from STAR*D that outcomes are relatively poor when compared to the OAD+PBO arm in the TRANSFORM-2 trial. [redacted] compared to STAR*D (13% remission and 16.3% response). This further suggests that keeping the treatment

effect constant for the OAD+PBO is conservative because we are maintaining a higher treatment effect than observed in STAR*D for the comparator arm.

There is a strong precedent of generalising earlier lines of data to later treatment lines in depression that comes from both the NICE TA367 (*vortioxetine for treating major depressive episodes)* [19] and from the current NICE clinical guideline 90: *Depression in adults: recognition and management* and the ongoing consultation of the new draft NICE guideline. Both the technology appraisal committee and guideline committee generalised earlier lines of data to later lines of treatment. In NICE TA367, the manufacturer used their 2nd line MDD efficacy and adjusted this to the 3rd line MDD population using data from the STAR*D study. The committee in TA367 used the manufacturer's analysis to make the recommendation for vortioxetine despite the company having no clinical evidence in that patient population. It is noted that the presence of SUSTAIN-2 data in the 3+ prior OAD and after augmentation population means that there is already additional evidence than what was available to the committee to make a positive recommendation in TA367.

In the current NICE clinical guideline 90: *Depression in adults: recognition and management* and the current updated version of the guideline that has recently been available for consultation, first line data have been used to inform the later line efficacy and recommendations for later line treatment options. Given the data limitations in this disease area, the guideline committee were required to make a number of extrapolations and assumptions for both psychological and pharmacological treatments from the evidence review for first line treatments to later line recommendations. Generalising the ESK-NS proportional treatment effect to one later line was also supported by the clinical experts consulted in preparing this addendum (Appendix B). Feedback from the clinical experts consulted during the advisory board noted that the [redacted]

The experts also cited a similar rationale to the clinical expert in 3rd ACM regarding the new mechanism of action of ESK-NS supporting at least a consistent treatment effect. In other words, as people continue to fail subsequent OADs they are likely to see a reduction of effect, whereas people who switch to a new mechanism of action, like ESK-NS, will see an increase in effect in terms of response and remission, and regardless of line of treatment given the benefit of a new mechanism of action. Given the increased relative treatment effect seen while moving between lines for ESK-NS in TRANSFORM-2, we reason that the approach to reduce the efficacy for ESK-NS based on SUSTAIN-2 data but maintain the treatment effect for the OAD+PBO arm is highly conservative. This assumption is further justified by the high number of people in the TRANSFORM-2 trial in the 3+ OAD failures subgroup, who had failed an augmentation treatment. In the trial, [redacted] of the patients in the 3+ OAD failures subgroup had also failed augmentation treatment and therefore were already a relative severe population.

We therefore believe the approach in the base case, which is to [redacted] but to maintain the efficacy of the comparator (from the 3+ failures TRANSFORM inputs) should be considered a highly conservative

estimate of the efficacy in the new proposed subpopulation (after 3+ OAD failures and after augmentation).

1.3 Monotherapy OAD remains the main comparator for ESK-NS in a 3+ OAD and after augmentation population with treatment effect for an OAD conservatively assumed as maintained between lines of treatment

Janssen consider OAD+PBO as the main comparator in the 3+ OAD failures and after augmentation population. Although we acknowledge that, in this population, other comparators may be used in clinical practice, a monotherapy OAD remains the most relevant comparator for ESK-NS for several reasons: firstly, [redacted] outcomes for other comparators included in the scope, 'such as combination or augmentation therapy and ECT, were highly uncertain.' The lack of evidence in the company NMA also highlights the lack of evidence for other comparators listed in the scope. Secondly, there is substantial use of the OADs (49%) and relatively limited use of augmentation strategies (17%) at this stage of the pathway in clinical practice. [20] Those who have tried augmentation strategies are also less likely to try them subsequently if they have not responded previously.

In addition to OAD monotherapy being the most used treatment at this stage in the treatment pathway, we consider that the comparison with OAD is likely to be conservative given that published meta-analyses have shown that ESK-NS had a relative effect size that was nearly twice as high versus antidepressant augmentation with second-generation antipsychotics than compared with a monotherapy OAD. [21] Furthermore, the same analysis showed that treatment with ESK-NS resulted in a mean reduction of Montgomery-Asberg Depression Rating Scale (MADRS) total score from baseline that was >7 points greater than augmentation with second-generation antipsychotics. [21] It is also important to note that the adverse effect profile of augmentation treatments may be unfavourable to some patients. [22] Lastly, using a cost of OAD in the economic model instead of the cost of an augmentation strategy is conservative, given augmentation treatments have higher drug costs and significantly higher monitoring costs, especially for treatments like lithium. [5]

We note in TA367 (vortioxetine) that the manufacturer adjusted the treatment effect for both the intervention and the comparator downwards, based on the STAR*D study. The section of the guidance is reproduced below:

> The original analyses used an absolute probability of remission for vortioxetine from the REVIVE trial. Relative treatment effects for the comparators would then be applied to this value. In order to reflect the fact that the decision problem is now considering 3rd line treatment, the manufacturer adjusted the absolute probabilities of response, remission and no-response on vortioxetine. The adjustment was based on a proportionate reduction of the REVIVE numbers based on the proportionate reduction from 2nd to 3rd line observed in STAR*D trial (Rush et al., 2006) – see Table 4 in response to ACD. The adjusted probability of a patient remitting after 3rd line with vortioxetine is now assumed at 18.1% (decreased from 40.5% in the original submission for 2nd line), and the probability of no response at 44.8% (increased from 38.5% in the original submission for 2nd line). Note that no adjustment needed to be made to the probability of response without remission since this was the residual probability after accounting for remission, no response and withdrawal due to AEs. Also note that the probability of withdrawing due to AEs was not assumed to differ between 2nd and 3rd line.

In contrast to TA367 which used data from the STAR*D trial to adjust estimates of later line efficacy, we have not changed the efficacy for the comparator OAD but have conservatively used SUSTAIN-2 data to adjust the ESK-NS efficacy.

Overall, Janssen believe that above approach is likely to be highly conservative. There is a strong precedent in NICE's decision making for assuming treatment effect is maintained at next line of therapy and supportive evidence from the TRANSFORM-2 trial that the relative benefit of ESK-NS increases with increasing line of therapy demonstrates that this is a strongly conservative assumption. Janssen recognise that although highly conservative there is uncertainty in the assumption and as such have explored alternative scenarios to support committee decision making (presented in Section 5) including no reduction in absolute treatment effect for ESK-NS, as well as decreasing the absolute treatment effect for the OAD.

1.4 The 3+ OAD failures subgroup efficacy is supported by the evidence from open label ESK-NS clinical trials and emerging RWE evidence

We note the Committee's concerns in ACM3, [redacted]

Janssen note the relatively small size of the subgroup in TRANSFORM-2 and have presented supportive evidence from the SUSTAIN-2 induction phase for this subgroup. [redacted] Table 5 compares the response and remission between the trials showing [redacted]

Table 5: Comparison of ESK-NS treatment effect at week 4 between TRANSFORM-2 and SUSTAIN 2 3+ OAD failures and after augmentation subgroups

	Remission	**Response**
TRANSFORM-2 (Non-response to at least 3 prior OADs)	[redacted]	[redacted]
SUSTAIN-2 induction phase (Non-response to at least 3 prior OADs)	[redacted]	[redacted]

This [redacted] response and remission rates [redacted] in the 3+ OAD subgroup has also been observed in an RWE cohort of patients using ESK-NS in France. The French RWE study included patients that had failed on average 3.7 previous treatments including augmentation treatments for most patients, see section 2.2 below. As highlighted in previous responses, we believe that the ESK-NS clinical trials are generalisable to UK clinical practice but accept that there are likely to be some variations between clinical practice and trials. With

regards to consistency of treatment effect differences between subgroups by line of treatment in TRANSFORM-1 and TRANSFORM-3. Firstly, it is important to note that TRANSFORM-1 is not an appropriate trial for consideration given that the dosing was not aligned with the licensed indication dosing recommendation. Data from TRANSFORM-3 was presented in *Janssen response to ACD 2* and were used to populate the scenario presented to the committee for the 3+ OAD failures subgroup.

1.5 The 3+ OAD failures and after augmentation population significantly improves previous implementation concerns given the smaller eligible population and due to patients already likely to be receiving care in a secondary mental health setting

Janssen response:

Janssen understands the Committee's concerns around a potential significant investment in ESK-NS, but as was noted in the survey of UK mental health trust pharmacists presented in the *Janssen response to ACD 2,* they were not able to identify any significant investment (apart from drug cabinets) that would be required and believed that existing facilities could be used within secondary care setting like ECT clinics. [23] The view of the pharmacists is aligned to those views [redacted] from clinical experts with respect to where ESK-NS would be used in the clinical pathway (i.e., after augmentation therapy). The new positioning of ESK-NS in the pathway after augmentation means that virtually all patients will be managed within a secondary care setting given that augmentation is not a primary care intervention. [5] If ESK-NS is used in this position in the pathway, then this supports the pharmacists view that implementation is manageable given many of the items and requirements are already in place to manage controlled drugs in a secondary mental health setting. Similarly, administering drugs that have relatively high treatment burdens like clozapine, antipsychotics and lithium and invasive treatments like electroconvulsive therapy (ECT) and transcranial magnetic stimulation (TMS) are also routinely carried out in this setting. Furthermore, this is also aligned to where infrastructure and space is currently located to administer ESK-NS based on the readiness survey that was presented to the committee in response to ACD 2. [24]

We respect the NHS commissioning experts' perspective in [redacted] and welcome the increased investment in community care and treatment closer to home for patients. We strongly believe these are important initiatives that will improve the care that people with depression receive. We are also conscious that there is a small but significant number of people who drive much of the treatment burden for the NHS in depression. This aligns to both the original TRD cost study and new analysis presented in Section 3, which show hospital admissions with a large length of stay and use of crisis resolution home treatment teams (CRHTT) is a significant cost for the NHS. Overall, this burden will likely decrease with the right targeted investment early in the pathway, as planned, but there will always remain a significant minority of patients, who will unfortunately continue to develop TRD, fail existing treatments, and will require intensive treatment in a secondary mental health care setting. The positioning of ESK-NS as an option after augmentation also helps to support those people with the highest burden of illness and those mostly likely to end up in hospital for their treatment of depression, see Section 1.1. This is where there is the highest unmet need for a new mechanism of action and where an intervention can have an important effect in breaking the cycle of TRD. Practically a community-based model of care would not support this positioning in the pathway given the small number of people who reach this stage of the treatment pathway, see

Appendix D. This also does not fit with the clinical experts' expected use of ESK-NS as outlined in [redacted] given the treatment burden. Janssen believe that ESK-NS could be a clinically appropriate cost-effective treatment for wider group of patients in time but believes some real-world evidence and clinical use is required before seeking a wider recommendation for ESK-NS and to support the investment case required to allow this to happen as the commissioning expert clearly identified would be the case.

Janssen believes proposing an additional treatment population later in the pathway (in 3+ OAD failures and after augmentation population) also significantly reduces the number of eligible patients for treatment and therefore makes implementation concerns more manageable for the NHS. Given the positioning after augmentation, the likely size of the total population pool has been estimated at 46,131 for the proposed 3+ OAD failures population and 14,745 to 15,940 at the proposed 3+ OAD failures and augmentation position, see Appendix D. ESK-NS use would only be a small percentage of this patient populations. This positioning after augmentation particularly allows for a managed introduction and use of ESK-NS in clinical practice.

[redacted]

This is because patients later in the TRD treatment pathway have few treatment options. The analogue would be similar to ECT, which is associated with inequity already due patients needing to travel to access it. Janssen consider this is aligned with current ongoing discussions with government regarding setting up regional centres who can administer and provide innovative emerging treatment options, which have been designed to target those patients with the highest unmet need. This approach also fits with the use of existing infrastructure like ECT services.

Overall, we take the commissioning experts concerns seriously and as outlined before we have committed to several measures to ensure the appropriate use of ESK-NS including a registry system. [25] The positioning of ESK-NS later in the pathway largely mitigates these concerns given the smaller number of eligible patients. Janssen have therefore not assumed any additional costs of implementation for the 3+ OAD failures and after augmentation population, because we consider these are largely mitigated for the reasons above. To account for some additional costs given the smaller population, we have included a 1:1 cost for the ratio of nurse: patient per each administration of ESK-NS in the 3+ OAD failures and after augmentation population.

For the 3+ OAD failures population, we acknowledge that there could be some implementation costs given the larger population size and a requirement to set up facilities in community care in time depending on uptake, but we are still currently unclear what these may consist of given the previous survey with mental health pharmacists and note that any implementation costs would be one-off. These remain an irreducible uncertainty for the 3+ OAD failures population and will depend on uptake and use of ESK-NS if recommended and these have not currently been considered in the economic model.

2. New evidence from the long term ESK-NS safety studies shows that ESK-NS has a manageable safety profile, while emerging RWE demonstrates a consistent efficacy profile to the esketamine RCTs.

Janssen response:

To address [redacted], Janssen would like to present new evidence to the committee on the use of ESK-NS from three data sources:

- The phase 3 clinical trials, SUSTAIN-2 and SUSTAIN-3.
- Two RWE studies from a French and Spanish cohort receiving ESK-NS in clinical practice.

2.1 Safety Data Update (SUSTAIN-2 and SUSTAIN-3): demonstrates that esketamine has a favourable safety profile with acceptable tolerability

Janssen have presented updated safety data from two long-term safety trials in people who received ESK-NS. The design of the trials and where patients were sourced from i.e., parent trials, or previous ESK-NS trials, varied. The first study, SUSTAIN-2 is a safety trial of patients with TRD treated with ESK-NS + OAD with repeated doses at intervals determined by symptom severity, while the second study, SUSTAIN-3 is an intermittently dosed study of ESK-NS + OAD in patients with TRD. A summary of the trials and conclusions is detailed below, and further information is discussed in Appendix E.

SUSTAIN-2: is a long term open-label safety study for people with TRD, which is being carried out in 21 countries and across 115 sites. Overall, long-term safety of intermittent treatment with ESK-NS + OAD was favourable with acceptable tolerability. The most clinically relevant safety findings were transient (i.e., resolving on the day of esketamine administration or within 1.5 hr of administration).

The long-term safety of intermittent ESK-NS administration (including patients ≥65 years of age) in TRD was favourable with acceptable tolerability following a 4-week induction phase of twice weekly dosing and up to 48 weeks of optimisation/maintenance phase with once a week or every other week dosing. [26]

SUSTAIN-3: is an ongoing phase 3 trial to confirm the long-term safety and efficacy of individualised, intermittently dosed ESK-NS + OAD in patients with TRD (defined as non-response to at least 2 different prior antidepressant treatments in the current depressive episode). The study is an open-label long-term extension to monitor repeated doses of ESK-NS at intervals determined by symptom severity. The primary objective of the study is to assess safety and tolerability.

Overall, there were no new safety concerns identified with continued intermittent ESK-NS dosing of up to 58 months in this study. Long-term exposure to ESK-NS resulted in no additional concerns or trends related to cognition, suicidality, events suggestive of abuse potential, lower urinary tract symptoms, and renal or hepatic adverse events. The study provides 3,034.2 patient-years of exposure to ESK-NS. Therefore, it can be concluded that the long-term safety of ESK-NS administration in patients with TRD (including those ≥65 years of age) was favourable with acceptable tolerability following cumulative exposure. [27]

In conclusion, both long-term studies highlight that the safety of esketamine nasal spray was favourable with an acceptable tolerability, and that long-term exposure to esketamine resulted in no additional safety concerns.

2.2 The French ESKALE study and Spanish compassionate use program provide supportive evidence in a real-world population regarding overall treatment efficacy for ESK-NS

Janssen response:

Since the last ACM, initial results of a French real-world evidence study (ESKALE) of ESK-NS for patients with TRD and initial results from a Spanish compassionate use programme of ESK-NS have become available [28]. These data, despite initial use in populations likely to be significantly more severe and treatment resistant than populations studied in the ESK-NS clinical trial program given they were mainly used in compassionate use programs, nevertheless, provide supporting information on the efficacy and safety of ESK-NS.

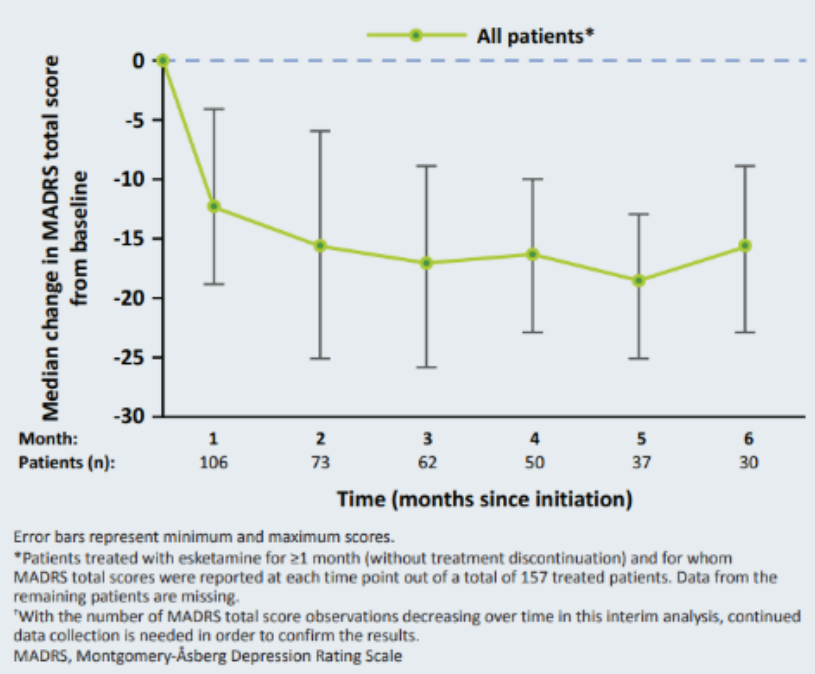
ESKALE study

In August 2019, four months prior to the European Marketing Authorisation for ESK-NS, the French authorities approved temporary use of ESK-NS (up to license) through the Autorisation Temporaire d'Utilisation (ATU) program. The ATU programme provides early access (pre license) to patients with a severe or rare disease with high unmet need and enabled RWE to be collected through the ESKALE study. Data was collected post-license and post launch. The key aim of the study was to collect clinical outcomes and key characteristics of people treated with ESK-NS in TRD. ESKALE was a retrospective, observational study of adults aged 18 years and over, with moderate-to-severe TRD, (where TRD is defined as non-response to ≥2 oral antidepressant drugs). We present the results and characteristics of a recent interim analysis of the study observed at 6 months of follow-up [29]

ESKALE includes 160 patients from across France. Patients were categorised into three different cohorts based on initiation into the study and if it was during 1) the temporary authorisation license period (early access), 2) post-license period or 3) post-launch period. There was a data lock in the study on 22nd September 2021, one year from the start of the temporary authorisation period. Overall, data was generated for 157 patients since 3 people dropped out during the temporary authorisation period. In total there were 14, 41, and 102 people recruited to the study in the temporary authorisation period, post-license period and post-launch period, respectively. [29] Baseline characteristics are presented in Appendix F.

After 4 months of treatment with ESK-NS, patients experienced a median decrease of 16.5 points, see Figure 1, from baseline when measuring MADRS, corresponding to 58.0% of responders and 43.1% of remitters. The results demonstrate the consistent clinical benefit that ESK-NS brings to real world patients, as indicated by the Phase 3 trial programme. Although the inclusion of patients into the study was defined as non-response to ≥2 oral antidepressant drugs, patients on average were receiving esketamine as 4th line treatment (3.7). The data at 4 months is therefore also supportive of the efficacy seen in the 3+ OAD failures subgroup from TRANSFORM-2.

Figure 1 Change in MADRS total score from baseline (N=157)



Error bars represent minimum and maximum scores.
*Patients treated with esketamine for ≥1 month (without treatment discontinuation) and for whom MADRS total scores were reported at each time point out of a total of 157 treated patients. Data from the remaining patients are missing.
†With the number of MADRS total score observations decreasing over time in this interim analysis, continued data collection is needed in order to confirm the results.
MADRS, Montgomery-Åsberg Depression Rating Scale

The results at month 6 have been reported for those patients who have reached that timepoint in the study, but with a lower number of observations (n=30), which consists of 14 patients who were in the original ATU cohort. The ATU programme provides early access (pre license) to patients with a severe or rare disease with high unmet need and therefore the initial cohort is likely to have been a cohort who were sicker and less reflective of a general TRD patient in clinical practice. Therefore, the current results at 6 months should be interpreted with caution until further data is collected. However, after 6 months of treatment with ESK-NS, patients experienced a decrease of 16.0 points in median MADRS total score from baseline, which corresponds to 46.7% of responders and 26.7% of remitters. These results can be seen in Figure 1. Furthermore, ESK-NS reduced the median Patient Health Questionnaire (PHQ-9) total score by 9.0 points, which was observed 4 months after esketamine initiation. Figures of the response and remitters reductions can be shown in Appendix F. The initial data collected provides sufficient supportive evidence of the clinical benefit of ESK-NS in a real-world population over a longer time, however data collection is ongoing which will inform future analyses.

Spanish ESK-NS compassionate use program

Similar to the ATU cohort in France, there are some initial results from a compassionate use program of ESK-NS in Spain. A compassionate use program was implemented for depressed patients after failing to respond to two or more proper antidepressant trials, one augmentation strategy and a non-pharmacological therapy e.g., ECT if accepted by the patient and not contraindicated. The aim of the study was to describe the effectiveness and tolerability of ESK-NS.

A total of 32 patients were included, with a mean age of 54.9 years and 69.9% who were females. ECT had been used in 46.9% (n=15) and of the remaining 53.1% (n=17) patients, 58.8% (n=10) rejected the treatment, 29.4% (n=5) did not have access to this procedure and in 11.8% (n=2) of cases ECT was partially contraindicated. This cohort was a significantly sicker cohort aligned with compassionate use. ESK-NS was effective in 87.5% (n=28) of patients with response and remission

rates after 6 months being 56.3% (n=18) and 31.3% (n=10), respectively. The majority of responders, 55.6% (n=10) responded during the first week and 22.2% (n=3) during the first month. Adverse events were mild, and tolerability was good with dizziness in 15.6% of patients (n=5), dissociative symptoms in 9.4% (n=3), anxiety in 3.1% (n=1) and 71.1% (n=23) reported no adverse effects).

Overall, both the ESKALE study and Spanish compassionate program, despite studying cohorts that were more severe than seen in the ESK-NS clinical trials and in the optimised populations proposed to NICE, show ESK-NS to be effective with a manageable tolerability similar to the clinical trials and supportive of the ESK-NS treatment effect being maintained at later treatment lines. Most importantly they show that for people who have failed multiple treatments, ESK-NS shows a fast onset and positive response and remission rates in people with TRD that have failed many previous treatment options.

3. **A new analysis of UK data demonstrates the high resource utilisation and high healthcare cost of treating people with TRD, supporting the original TRD cost study rather than the Byford et al study, which is not appropriate given the setting of care.**

Janssen response:

3.1 Using the Byford et al study solely significantly underestimates the NHS cost of treating patients with TRD

Healthcare resource utilisation and costs of the health states is a key driver of the economic model. The uncertainty that the committee have identified in [redacted] is caused by the divergence in the two sources used in the economic modelling, the Byford et al study, [30] and the TRD cost study subsequently published as Denee et al, [11]. Janssen disagree that the Byford et al study is an appropriate source to be used solely in the economic modelling, since this study was conducted in a primary care population and is likely to significantly underestimate the total and average cost of treating a person with TRD. As noted below, especially considering the later line positioning in the new proposed subgroup (3+ OAD failures and after augmentation treatment), the Byford study is not an appropriate study to inform the costs of these patients. The Byford et al study does not meet the definition of patients with TRD, as the patients only needed to be in receipt of 2 antidepressant medications and therefore had not failed 2 antidepressants.

In addition, it is important to emphasise the incomplete healthcare resource utilisation data related to TRD treatment within the General Practice Research Database (GPRD) database for the Byford et al study, which were captured and represented in the TRD cost study. This was the primary reason that Janssen decided to utilise the retrospective chart review design rather than a primary care database like GPRD to estimate the resource use for TRD patients in the UK. The linkage of GPRD to secondary care is notoriously poor, in part driven by the fact that while computerisation of primary care resource in UK general practice is very good, but handwritten records remain common in secondary care. [31] As acknowledged in the Byford et al paper, GPRD is significantly limited in its ability to represent secondary care resource use, namely consultations with specialists, and the granular detail of hospitalisations, length of stay and type of secondary care mental health contacts. The richness of the health data for secondary care contacts captured in the TRD cost study is clearly seen in Table 6. Without the individual specific psychiatric consultations and treatment sessions captured, the associated secondary care cost is likely to be significantly underestimated.

The biggest driver of the difference between the two studies is the variation in hospital costs, with an approximate annual comparison of £154 in Byford et al [30] for 'non-remitters' compared to £4,942 for patients in MDE in the TRD study. In Byford et al, hospitalisation data in GPRD was significantly limited (with no hospital length of stay data or psychiatric/intensive care unit admission data available), while CRHTT resource use were not captured at all. It is also important to note that just 6% of "non remitter" patients in the Byford et al sample were hospitalised versus 22% of patients in the MDE health state in the TRD cost study. This substantial difference suggests the GPRD database significantly under-reported the number of hospitalisations occurring within the sample, a known limitation of primary care databases such as GPRD and likely reflective of less severe and non-TRD cohort of patients. Hospitalisation length of stay information, known to be a significant cost

resource, is not available within GPRD, and instead an average cost was applied per admission (based on personal social services research unit (PSSRU) and hospital episode statistics (HES) data). This will have a disproportionately high impact on the 'non-remitters' healthcare resource use (HCRU) cost given the increased likelihood of hospitalisations occurring within this group. Byford et al. acknowledge the limitations associated with using an average hospitalisation duration (based on HES data) likely under-representing the hospitalisation resource used within their population. Comparisons with UK randomised controlled trials, such as the THREAD (THREshold for AntiDepressant response) [32] study, emphasise the notably lower hospitalisation resource reported by Byford et al.

Looking further into these studies, Byford applied a single cost (£1,887.60) applied to all hospitalisation admissions, with no granularity regarding duration or resources used during the hospitalisation. In comparison, the TRD cost study applied costs to inpatient hospitalisation (£266.12 per night without psychiatric ward admission, £404 with psychiatric ward admission) and admission to intensive care units (ICU) (£1,328 per night). The TRD cost study captured the total number of hospitalisations for each patient, followed by a detailed collection of data for each hospitalisation (date of admission and discharge, admission via the emergency room (ER), admission to psychiatric ward and any days spent in ICU). All of the details captured were factored into the derived hospitalisation costs, with particularly high unit costs associated with admission to the psychiatric ward (£404 per night) or ICU (£1,328 per night).

Table 6: **Comparison of resource categories included in the Byford et al and TRD costs study**

Resources categories included in Byford et al, 2011	Resources categories included in TRD cost study, Denee et al, 2021
Antidepressant use	
Antidepressant use	Antidepressant therapy
Concomitant medication	Concomitant medication
Primary care contacts	
GP visits	GP visits
GP phone calls	Nurse visits
Secondary care contacts	
A&E	A&E
Inpatient days (no length of stay data)	ICU (length of stay data)
Other specialist contacts	Inpatient Psychiatric ward (length of stay data)
Psychiatrist contacts	Inpatient Non-psychiatrist ward (length of stay data)
Psychotherapy	CRHTT visits
	Specialist visits
	Occupational therapy
	CBT
	Counselling
	Psychotherapy
	Mindfulness therapy
	ECT
	TMS
	Health coaching
	Behavioural activation therapy

Additionally, the Byford et al study did not capture CRHTT use or costs, likely a limitation of GPRD, and therefore this resource is not represented within the total derived costs. Failure to consider CRHTT support within the patient pathway will therefore inevitably result in under-representation of the 'true' cost of managing TRD. The TRD cost study captured use of CRHTTs. CRHTT is an important alternative to inpatient hospital care for service users with serious mental illness, offering flexible, home-based care, 24 hours a day, seven days a week. The main target group will usually be adults between 16-65 years of age, whose mental health problems are of such severity that they are at risk of requiring psychiatric hospitalisation. It is also important to note that use of CRHTTs has increased significantly in the last 10 years and therefore a limitation of the Byford et al study given when the analysis was conducted and published.

Furthermore, the use of a general retrospective database is a significant limitation of the Byford study compared to the TRD cost study. The use of a retrospective chart review study design in the TRD cost study meant that it was possible to determine health state changes within the follow-up window for each patient, and therefore allocated healthcare resources to each health state. In comparison, Byford et al were unable to identify health state changes within the follow-up window for each patient and were limited to categorising patients as either 'remitters' or 'non-remitters' based on their antidepressant prescription history, failing to account for the fact that fluctuations in patient health states will have occurred. The derived 'non-remitter' healthcare resource utilisation (HCRU) cost by Byford et al will therefore include resources used across various health states (albeit in patients ultimately classified as 'non-remitters'), and therefore comparison with the MDE costs derived in the TRD cost study should be made with considerable caution. Overall, we believe that the Byford et al study has serious limitations when generalising to a population with TRD and although maybe adequate to estimate resource use for people with non-TRD MDD in primary care, is likely to significantly underestimate costs of secondary care mental health costs and people with TRD.

3.2 The significant NHS healthcare resources that patients with TRD utilise is highlighted in the TRD cost study and is further confirmed in a recently conducted retrospective database study

It is important to note the revised position later in the pathway is likely to mean that the costs of health states in the model will be more reflective of the TRD cost study, as a 3+ OAD failures and after augmentation population will be a population that is largely managed in a secondary care mental health setting given that augmentation treatments are primarily initiated by a psychiatrist. In addition, resource use and costs are likely to increase by line of therapy, as has been discussed in the Section 1.1.

In addition to the study that was presented in Section 1.1 where the significant resource use for people with TRD from Northwest London was discussed. Janssen presents an additional source of data that has been conducted in a TRD population in a secondary care mental health setting using the Clinical Record Interactive Search (CRIS) database at South London and Maudsley, NHS Foundation Trust, Appendix G. These data were collected from a TRD population in a secondary mental health setting. The following healthcare contacts were extracted and analysed and then costed:

- Inpatient bed nights
- Mental health community contacts
- Mental Health home treatment teams
- Emergency contacts at A&E and crisis contacts
- ECT and CBT contacts
- Face to face and remote contacts for clinicians, nurses, occupational therapy, psychotherapy, Social work.

In the primary analysis, [redacted] met the definition of TRD of at least 2 prior OAD failures. [redacted]. The burden of illness was high [redacted] inpatient bed days [redacted] The average length of stay was [redacted]. Overall, there were [redacted] face-to-face contacts by the cohort in the 6-month period or an average of [redacted] face-to-face contacts. There were on average [redacted] contacts with accident and emergency (A&E) services. The burden [redacted] for the number of previous episodes recorded with people [redacted] antidepressants having a cost of [redacted] within a 6-month period or [redacted] in a 28-day period. [redacted], which means that it is possible that patients who were no longer in their depressive episode (who are likely to use less HCRU) [redacted]

[redacted]

3.3 Janssen has provided a new scenario that uses the average of the Byford et al and TRD cost study

The TRD cost study is the most appropriate source of costs for people who have failed a number of previous treatment options and aligned to where ESK-NS would be used in clinical practice. This is supported by the CRIS database analysis above using a different methodology to the TRD cost study, but which shows consistent results. We have therefore continued to include the TRD cost study in the base case analysis of the economic model for people who have failed 3+ OAD failures and after augmentation. This is especially pertinent for those people who have failed augmentation treatment and who will be managed by psychiatrists in a secondary care setting. These are the small number of people that use most healthcare resource, as these are the most resistant and severe patients who have failed all other treatment options. Therefore, it is more appropriate to use the TRD study in the model to capture the likely resources that are used by these people.

For the 3+ OAD failures group, Janssen note the committee's previous conclusion that the costs are likely to lie in between the company's and the ERG's approaches. Above, we have provided further

rationale for why the Byford study is less appropriate. For the people who have failed 3+ OADs we used a weighted average of the Byford and the TRD cost study to reflect the Committee's comments. This includes a 25:75 average of the Byford et al and TRD costs study given the limitations in generalisability associated with the Byford et al study. This leads to an average cost of £824.00 (inflated) for the MDE health state in the model.

As noted above, for health state costs used in the model, we have inflated the costs to 2020/2021 prices (Appendix J). Since the ERG inflated Byford study costs at the time of the 3rd ACM (2020), we have further inflated these by 1 year to bring them in line with TRD study inflated costs.

4. The ERG treatment cap in the model and subsequent treatment efficacy is overestimating long-term outcomes for people with TRD based on the literature

Janssen response:

4.1 The current ERG treatment cap based on Wu et al has been corrected to increase face validity, but remains a highly conservative estimate of subsequent treatment efficacy

The ERG cap implemented on relapse and loss of response boosted the efficacy of subsequent treatments in the economic model and in doing so reduced the number of people in the MDE health state overtime. The company model outputs suggested that 76% of the disease course was spent in the MDE health state reflecting an observed low remission and a high rate of relapse and loss of response seen in the literature [12]. Janssen subsequently provided a targeted literature review which also demonstrated a low longer-term remission rate for people with TRD from a number of studies, in particular, a couple of UK TRD studies which showed a remission rate of 6.5% at 18 months decreasing to 4.4% after 42 months. [33] Additionally, the Sackheim et al [34] findings where there was only a 28.90% chance of remaining well for 12 months after acute remission from the STAR*D study at Step 4 (failure after 3 lines of therapy). Sackheim, [34] also report only a 3.76% probability of sustained benefit at Level 4 (step 4). Overall, Janssen agrees with the committee that there was significant heterogeneity in the definition of remission, response, and trial designs of the trials included in the targeted literature review. However, the long-term outcomes remain an inherent and irreducible uncertainty in the literature of the disease area and is reflective of the limited research that has been done in depression and especially in people with TRD, which should not unfairly bias against ESK-NS. Janssen have presented additional new evidence in Section 4.2 that reinforce the poor outcomes people with TRD face and suggest that the current ERG cap is underestimating the level of relapse in the model for subsequent treatments.

The Wu et al. study [35] used by the ERG to support the treatment cap, and as noted previously in our response to the *ERG's critique of Janssen's response to ACD2*, is likely to have low generalisability to a UK health care setting given the difference between the way that mental health is treated in the US compared to the UK. The study showed that the mean length of the first TRD episode was 1.56 years, and the mean length of remission was 0.90 years. The clinical expert noted at the 3rd ACM that this is likely to be optimistic and commented that he felt the true estimate may lie between the ERG and Janssen's assumptions. In addition, the model results when using the ERG cap, derived from Wu et al study lack face validity currently, as they lead to a lower relapse for subsequent treatments than the 3rd line comparator OAD relapse rate in the model. It is generally accepted that relapse rates increase with each additional line of therapy. [12] The current Wu et al study and the ERG cap therefore leads to a conservative estimate of the relapse rates when compared to the literature and compared to recent literature published in the next section.

4.2 In additional to previous evidence submitted ahead of the 3rd ACM, new evidence suggests that a proportion of TRD patients are likely to spend a significant period of time in MDE due to low levels of remission and high levels of relapse

A UK study using the DISCOVER dataset, which was described in Section 1.1, has estimated the length of a depression episode for people with MDD and TRD in the UK. The results show that the depressive episode is much longer and there is a high proportion of people who relapse at later lines of treatment. This is supportive of a higher relapse rate for subsequent treatments used in the economic model. As shown in Table 7 below the average MDD duration was [redacted] and for TRD patients it was [redacted] This data suggests that people with both MDD and TRD in the UK spend [redacted] than the 1.56 years demonstrated in the Wu et al study.

Table 7: DISCOVER dataset average duration of depression

Average depression duration in months	MDD patients	TRD patients (≥2 lines of therapy)	TRD patients (≥3 lines of therapy)	TRD patients (≥4 lines of therapy)
Mean	[redacted]	[redacted]	[redacted]	[redacted]
Standard Error	[redacted]	[redacted]	[redacted]	[redacted]
Median	[redacted]	[redacted]	[redacted]	[redacted]
Standard Deviation	[redacted]	[redacted]	[redacted]	[redacted]

In addition, the study looked at the treatment response for patients over a [redacted] days after the end of the previous treatment sequence, in which the previous treatment sequence may have resulted in [redacted]. Remission was defined [redacted] following the previous treatment sequence or a [redacted] There was a difference in [redacted] The study shows [redacted] within the TRD cohort.

Table 8: Relapse and remission seen in a UK cohort

MDD or TRD definition	Total cohort size	Number of instances	Number of patients	% of cohort
MDD group - Relapse	[redacted]	[redacted]	[redacted]	[redacted]
MDD group - Remission	[redacted]	[redacted]	[redacted]	[redacted]
TRD group - Relapse	[redacted]	[redacted]	[redacted]	[redacted]
TRD group - Remission	[redacted]	[redacted]	[redacted]	[redacted]

In addition, a separate independent study has been published in 178 TRD patients from 3 UK centres that is supportive of UK TRD patients experiencing longer episodes than seen in the Wu et al. study with the median duration of 5 years (95% CI 4-6 years). [35]

Similarly, a new study looking at the real-world evidence from a European cohort of patients TRD (including patients from the UK) shows that among 441 patients enrolled, after 6-months post-initiation of a new treatment strategy, as per routine clinical practice, only 16.7% achieved remission and 73.5% showed no response [36] . In addition to having poor health-related quality of life (HRQoL) and reduced function. At month 12, while 19.2% achieved remission and 69.2% showed no response, 33% of those in remission at month 6 were no longer in remission. Most worryingly, at month 12 despite the poor outcomes, 60% of patients had not changed treatment since enrolment.

The UK sub cohort of the above European study confirms the characteristics of the TRD episode duration and low remission rates are from across the UK (n=49). The analysis of the UK cohort illustrated the mean duration of their current episode to be 6.1 years, and a remission rate which was lower than the rest of European cohort with only 8.9% of people being in remission after 6 months of treatment and a further 8.9% having response without remission. [37] This data further supports that the available evidence identified by the literature review conducted (included in the ACD 2 response) by the company and suggest far worse treatment outcomes for patients with TRD in the UK than seen in the US Wu et al study.

4.3 The ERG cap has been corrected to maintain face validity in the model, and the amended ERG cap should be considered the upper limit of the cost effectiveness estimate given the literature

To provide an explanation of our approach for subsequent treatment efficacy in both submitted models, it is important to provide some background. The STAR*D study reports results for Steps 1 through to Step 4. From Step 5 (which correlates with after 4 treatments failures), we originally used the ratio between Step 3 and Step 4 to extrapolate the efficacy (including relapse and loss of response). This was then repeated for each subsequent line based on the extrapolated data. The result is that relapse and loss of response increases as the STAR*D data is extrapolated into later subsequent treatment lines. The extrapolated relapse and loss of response for Step 5 and Step 6 can be found below:

Figure 2 Relapse risk for subsequent treatments, informed from STAR*D

4-week relapse risk from StarD	
6.77%	Exponential fit to Step 3 in Figure 3 from Rush et al.
12.79%	Exponential fit to Step 4 in Figure 3 from Rush et al.
24.18%	Step 5, based on extrapolation
45.70%	Step 6, based on extrapolation

Figure 3 Loss of response for subsequent treatments, informed from STAR*D

4-week loss of response risk from StarD	
22.16%	Exponential fit to Step 3 in Figure 4 from Rush et al.
22.81%	Exponential fit to Step 4 in Figure 4 from Rush et al.
23.49%	Step 5, based on extrapolation
24.18%	Step 6, based on extrapolation

Given STAR*D provides data to Step 4 only, the extrapolated data can be found in the shaded cells in Figure 2 and Figure 3. Reviewing the steps in the extrapolated data from STAR*D, as noted by the

ERG previously, the relapse input would exceed 100% after Step 8 which is not clinically plausible, and hence Janssen previously decided to cap this at 99%. Prior to the 3rd ACM, the ERG capped the relapse and loss of response for all subsequent treatments and non-specific mix using the STAR*D Step 4 value only (at 12.8% relapse and 22.8% loss of response), as can be seen below:

Table 9: ERG capped relapse and loss of response

ERG cap approach, as used in scenarios presented at 3rd ACM	Relapse risk	Loss of response risk
Comparator OAD relapse risk (at TRD line 2)	16.8%	23.1%
Subsequent treatments		
TRD line 3	12.8%	22.8%
TRD line 4	12.8%	22.8%
TRD line 5	12.8%	22.8%
BSC/ non-specific treatment mix	12.8%	22.8%

As can be seen Table 9 above, the treatment efficacy for subsequent therapies (12.8%) are lower than the relapse rate for the comparator OAD (16.8%). This is not clinically valid (i.e., this implies that patients relapse quicker at the earlier line of treatment in the model).

To explain this lack of face validity, the choice of the value of the cap neglected the fact that we calculated a weighted average for relapse and loss of response for the initial OAD (i.e., using step 4 and step 5 values of STAR*D), rather than using Step 4 values only. The method for this is presented in the Appendix H. We have therefore amended the cap, so the cap is at least consistent with the comparator OAD i.e., relapse rate is capped at 16.8%. We have provided a scenario (scenario 9) using the corrected ERG cap for loss of response and relapse in section 6.

In the base case, we have provided an alternative to the amended the ERG cap given the uncertainty and difference in estimates from the literature, and moved the cap to one line later as per the STAR*D extrapolations, as:

1) This amendment corrects the face validity of the previous ERG approach. Using the original ERG cap on subsequent treatment relapse and loss of response resulted in a lower relapse rate for subsequent treatments than previous treatments in the model.
2) Rather than setting the cap from Step 4, the cap should correlate with data from Step 5 of STAR*D data, as this is aligned with the first subsequent treatment in the economic model (i.e., 4 prior failed treatments).
3) The resulting output of the model when using the amended cap is supported by the evidence available in the longer-term studies (see section 4.2) suggesting that remission is lower, and relapse is higher in clinical practice.

Table 10 below presents a summary of the approach used for the subsequent treatments used in the models.

Table 10: Summary of subsequent treatment approach used in base case models

	Revised ERG cap for sub population 1: 3+ OAD failures	Approach for sub population 2 (at least 3+ OAD failures and after augmentation)
Relapse inputs		
TRD line 3 (4 prior failures)	0.318	N/A due to positioning
TRD line 4	0.318	0.318
TRD line 5	0.318	0.318
TRD line 6	N/A due to positioning	0.318
BSC/ non-specific treatment mix	0.318	0.318
Loss of response inputs		
TRD line 3	0.237	N/A due to positioning
TRD line 4	0.237	0.237
TRD line 5	0.237	0.237
TRD line 6	N/A due to positioning	0.237
BSC/ non-specific treatment mix	0.237	0.237

Overall, the result of introducing this correction to the ERG cap and increasing the ERG cap based on one later line has the impact of improving ICERs slightly while maintaining the face validity of inputs. Consistent with the new evidence presented above, the Markov Trace output predicted from the amended cap on subsequent efficacy from economic model (see section 4.3 below) is presented below. This shows the proportion of time that the 3+ OAD failures model spent in remission over the full-time horizon of the model:

Table 11: Markov trace outputs

Proportion of time spent in remission health state over economic model (OAD arm)	3+ prior OAD failures	3+ OAD failures and after augmentation
With the ERG cap on efficacy of subsequent treatments, as presented at the 3rd ACM	39.3%	-
Revised company approach (See section 4.3 below for methodology)	21.8%*	20.9%*

*Based on undiscounted proportion of Life Years spent in remission in the full model

It is important to note the consistency in remission results in the revised company approach from the available long-term TRD studies, although the remission rate is still slightly higher than many of the studies. Implementing the ERG cap (as per 3rd ACM) is a conservative estimate of the long-term remission outcomes of patients with TRD with nearly 40% of people in the model being in remission over time. The revised company approach to the ERG cap has therefore been used in the base case of the economic model, as this reflects many of the studies that have a lower rate of remission over time including the new evidence identified since the last appraisal committee and corrects the face validity of the inputs. We note that the capping approach of the long-term efficacy of the

subsequent treatments remains a conservative approach, given that later lines of treatment are associated with increasing rates of relapse.

5. **A new base case is included for the committee's consideration: overall, the new evidence and updated model shows that ESK-NS is a cost-effective option in both the 3+ OAD failures, and the 3+ OAD failures and after augmentation position in the MDD pathway. The incremental cost effectiveness ratios (ICERs) with the new value proposition are below £10,000 per QALY threshold**

Janssen want to support the committee with decision making options for ESK-NS access in England and Wales, and therefore have provided the committee with two updated base case models in two sub populations and several key scenario analyses to facilitate decision making. We have revised the cost effectiveness estimates for the sub population 1 (3+ OAD failures) and sub population 2 (3+ OAD failures and after augmentation) based on revised or new data, as outlined throughout this document. Janssen have ensured to maintain robustness and clinical plausibility with modelling assumptions and have made use of updated data and clinical expert opinion. In Appendix I, we outline changes that have been made to the company's base case, and in section 5.2 we explore key scenario analyses and the impact the scenarios have on the ICER.

We would like to remind the committee that a number of the committee's preferred assumptions used in the 3rd ACM cost effectiveness model were conservative in nature based on the data submitted. A reminder of these conservative assumption has been presented in Appendix I. For example, nurse monitoring, short-term and subsequent treatment efficacy, and the percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons. Overall, the model revisions should reduce the overall level of uncertainty in the appraisal and improve the Committee's confidence that ESK-NS is a cost-effective treatment option for people with later line TRD. Furthermore, we would like to remind the committee that Janssen have further significantly [redacted]

5.1 Revised base case economic model inputs

Janssen have outlined in Table 12 the key inputs for the two revised base case economic models, in addition to the updated incremental cost effectiveness ratios for each base case. Column 3 includes the key inputs for sub population 1 (3+ OAD failures population model) and column 4 includes key inputs for sub population 2 (the 3+ OAD failures and after augmentation population model). Column 2 includes the parameters that were included in the 3+ OAD failures population for the 3rd ACM.

Base Case sub population 1: 3+ OAD failures population

Prior to the 3rd ACM, Janssen submitted an economic model for the 3+ OAD failures population, which we would like to resubmit for consideration by the Committee at the 4th ACM, with some minor amendments. Since the Committee and ERG have reviewed this model previously prior to the 3rd ACM, the majority of inputs in the model have not been changed. However, two notable parameters which have been updated for this population include:

1) the ERG cap applied to subsequent treatments, with a correction/amendment (as noted in Section 4) and;
2) Given the committee's conclusion that the health state costs lie between Byford et al and the TRD cost study, the source for health state costs from a weighted average of the TRD cost study and Byford costs

Base case 2 – 3+ OAD failures and after augmentation population:

As noted, Janssen would like to submit a 2nd economic model base case for sub population 2, which is the population 1 line later than previously submitted. Since the committee and ERG have not reviewed this model before, some inputs are new. The majority of the inputs are aligned to the sub population 1 (3+ OAD failures model) base case, with revisions for the later line population, which includes:

1) the ERG cap applied to subsequent treatments, with a correction/amendment (see section 4)
2) short-term efficacy from TRANSFORM-2 adjusted for a population that have failed 3+ OAD failures and augmentation sourced from the SUSTAIN-2 trial induction period
3) dosing schedule increased to reflect the later line population where devices per session is increased based on extrapolation (see below)
4) administration costs for nurse monitoring per patient revised to 1:1 to reflect reduced numbers of patients with the smaller population
5) the source for health state costs is wholly from the TRD costing study and;
6) Percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons is decreased (from the 3+ OAD failures population) to 60% (see Appendix I)

See Appendix I for additional discussion on additional model inputs.

Table 12: Base case model inputs for 4th and 5th line

Model Input	3rd ACM sub population: 3+ OAD failures subgroup	4th ACM Sub Population 1: 3+ OAD failures subgroup	4th ACM Sub Population 2: 3+ OAD failures and augmentation for 4th ACM
Comparators	OAD	No change: OAD	No change: OAD
Time Horizon	20 years	No change: 20 years	No change: 20 years
Utility	TRANSFORM-2 and TRANSFORM-3	No change: In line with 3rd ACM model	No change: in line with 3rd ACM model
Carer Disutility	ERG method of incorporating carer disutility	No change: carer disutility included	No change: carer disutility included
Subsequent Treatment Efficacy and Long-Term Outcomes: Non-Specific Treatment Mix	ERG implemented cap on subsequent treatment efficacy (relapse and loss of response), based on Step 4 of STAR*D.	Revised ERG method: cap on relapse and loss of response based on weighted average of STAR*D Step 4 and Step 5 and moved one line later.	Revised ERG method: cap on relapse and loss of response based on weighted average of STAR*D Step 4 and Step 5 and moved one line later.
Short-term efficacy: updated source (SUSTAIN-2)	3+ prior failures subgroup data from TRANSFORM-2 and TRANSFORM-3	No change: In line with 3rd ACM model	Input source update: • ESK-NS efficacy reduced by proportional reduction from SUSTAIN-2 induction phase applied to TRANSFORM 3+ prior failures subgroup • OAD efficacy: No change, informed by TRANSFORM 3+ prior failures subgroup
Dosing Schedule	3+ prior failures subgroup data from ESK-NS clinical trials	No change: In line with 3rd ACM model	Dosing schedule update: • Treatment sessions remains • Increase in devices per session (extrapolated based on ITT and 3+ OAD failures data)
Administration Costs	Nurse: Patient Ratio 1:2 (£50.92 per administration)	No change: Nurse: Patient Ratio 1:2 (£50.92 per administration)	Revision based higher monitoring: Nurse: Patient Ratio1:1 (£72.92 per administration)
Health State Costs	Cost source: • TRD costing study • Byford et al study	Cost source revised: 75% TRD costing study & 25% Byford	Cost source: TRD costing study

Model Input	3rd ACM sub population: 3+ OAD failures subgroup	4th ACM Sub Population 1: 3+ OAD failures subgroup	4th ACM Sub Population 2: 3+ OAD failures and augmentation for 4th ACM
Costs Inflation	• TRD cost study not inflated • ERG inflated Byford to 2020 costs.	• TRD costs inflated to 2020/2021 • Byford costs inflated 1 additional year to 2020/2021	• TRD costs inflated to 2020/2021 • Byford costs inflated 1 additional year to 2020/2021
Percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons	Committee preferred: 70% discontinuation	Committee preferred: 70% discontinuation	Revision: 60% discontinuation in line with later line population (see Appendix I)
Adverse Events	As per TRANSFORM-2 data	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model
Mortality	No excess mortality for MDE health state	No change: No excess mortality	No change: No excess mortality
ICER (£/QALY)	[redacted]	[redacted]	[redacted]

5.2 Results: ICERs and model scenarios for revised subgroup for Committee decision making

In addition to the base case models, Janssen have presented several key univariate scenario analyses to test the variation in the ICER results for the 2 base case populations. The scenarios below are a brief description of key parameter changes for each base case and a rationale why each parameter has been revised.

Scenario analysis: 3+ OAD treatment failures population scenarios

- **Scenario 1 -** carer disutility excluded: during the 3rd ACM, the committee discussed that they would like to understand the impact on the ICERs if carer disutility was excluded from the base case model.
- **Scenario 2** - the percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons set to 60%: the committee agreed that an appropriate level of discontinuation in recovery was 70% for the 3+ OAD failures population. However, to allow for a more conservative assumption and review of ICERs we have further reduced this rate by 10% to 60%.
- **Scenario 3** - the time horizon has been reduced to a 5-year timeframe: the shorter time horizon of 5 years is consistent with all other depression models and a longer time horizon of 20 years (which is now the base case assumption) introduces uncertainty on modelling of future episodes, which are inherently associated with irreducible uncertainty. Furthermore, the cost effectiveness of ESK-NS in one episode is likely to be a proxy for cost effectiveness in future episodes.
- **Scenario 4** - a weighted average of costs inputs based on a 50% TRD cost study and 50% Byford study: although the more appropriate costing source for the health states in the late line populations is the TRD costing study, we have presented an extreme scenario of health state costs based on an even weight from Byford, which underrepresents the cost of later line secondary care mental health patients.
- **Scenario 5** - administration based on a nurse-to-patient monitoring ratio of 1:1: we have included the most conservative assumption of nurse-to-patient monitoring for this population, which is unlikely, and as noted by the clinical expert in the 3rd ACM would only "*be necessary when a service first starts administering esketamine, but that the ratio may increase to 1 nurse to a group of patients once the service becomes experienced and established*". However, as the most conservative monitoring ratio we have included as a scenario.
- **Scenario 6** - an assumption of Cuijpers et al as a source for excess mortality: in the recent NICE guideline for depression model, the source was used to have excess mortality associated with the MDE health state. [5]
- **Scenario 7** - 1.5% discount on costs and health effects**:** in line with the NICE reference case and due to the 20-year time horizon of the base case we are presenting a scenario of 1.5%.
- **Scenario 8** – increasing dosing of devices per session: all patients receive the maximum 84mg (3 devices) at every visit. We have provided a highly conservative assumption that all patients would receive the maximum ESK-NS dosage (number of devices) at each administration session.
- **Scenario 9**: Using the amended ERG cap (i.e., using 16.8% as relapse cap instead of original 12.8% relapse cap), but capping relapse and loss of response to 1st line model inputs as was previously done by the ERG.

Table 13: 3+ OAD failures population scenarios

Model Input	Scenario 1: No carer disutility	Scenario 2: Discontinuation reduced by 10% to 60%	Scenario 3: Shorter time horizon (5 years)	Scenario 4: 50/50 weighted Byford/TRD cost study	Scenario 5: nurse monitoring 1:1	Scenario 6: Excess mortality Cuijpers et al	Scenario 7: 1.5% discount on costs & benefits	Scenario 8: dosing max for devices and sessions	Scenario 9: ERG cap corrected for face validity
Comparators	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC
Time Horizon	20 years	20 years	**5 years**	20 years	20 years	20 years	20 years	20 years	20 years
Discount rate	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	**1.5% (for costs and health)**	3.5% (for costs and health)	3.5% (for costs and health)
Utility	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model	No change: In line with 3rd ACM model
Carer Disutility	**Carer disutility excluded**	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included
Subsequent Treatment Efficacy & LT Outcomes	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	ERG cap (revised) to 1st line cap	**ERG cap (revised) with 1st line model inputs (16.8% relapse, 23.1% loss of response)**
Dosing Schedule	Same as submitted for ACM 3	Same as submitted for ACM 3	Same as submitted for ACM 3	Same as submitted for ACM 3	Same as submitted for ACM 3	Same as submitted for ACM 3	Same as submitted for ACM 3	**Maximum devices* per session**	Same as submitted for ACM 3
Administration Costs	Nurse:Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)	**Nurse: Patient 1:1 (£72.92)**	Nurse: Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)	Nurse:Patient 1:2 (£50.92)
Health State Costs	75% TRD/ 25% Byford	75% TRD/ 25% Byford	75% TRD/ 25% Byford	**50% TRD/ 50% Byford**	75% TRD/ 25% Byford	75% TRD/ 25% Byford	75% TRD/ 25% Byford	75% TRD/ 25% Byford	75% TRD/ 25% Byford
Non-efficacy discontinuation of patients in recovery by 2 years	70%	**60%**	70%	70%	70%	70%	70%	70%	70%
Excess Mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	**Cuijpers (1.52 for MDE health state)**	No excess mortality	No excess mortality	No excess mortality
ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*For this scenario, the number of devices in Weeks 1-4 used an average of 2.875 devices, and all other weeks was 3.0 devices. This is reflective of the maximum average number of devices as per label wording (first dose should be 56mg, 2 devices).

Scenario analysis: 3+ OAD failures & after augmentation population scenarios

- **Scenario 1:** Carer disutility excluded: during the 3rd ACM, the Committee discussed that they would like to understand the impact on the ICERs if carer disutility was excluded from the base case model.
- **Scenario 2**: The percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons set to 50%: in line with the 3+ OAD failures scenario analysis, this scenario has reduced the discontinuation a further 10% to 50% to present conservative ICERs to the Committee.
- **Scenario 3**: The time horizon has been reduced to a 5-year timeframe
- **Scenario 4**: A weighted average of costs inputs based on a 75% TRD cost study and 25% Byford study: although the more appropriate costing source for the health states in the late line populations is the TRD costing study only, we have presented a conservative scenario for health state costs based, which includes 25% weight from Byford study.
- **Scenario 5**: Administration based on a nurse-to-patient monitoring ratio of 2:1: we have included a less conservative assumption of nurse-to-patient monitoring for this population, which is still likely conservative, but is comparative to the 3+ OAD failures base case
- **Scenario 6**: An assumption of Cuijpers et al as source for excess mortality: in the NICE guideline for depression, the source was used to have excess mortality associated with the MDE health state. [5]
- **Scenario 7:** 1.5% discount on costs and health effects**:** in line with the NICE reference case and due to the 20-year time horizon of the base case we are presenting a scenario of 1.5%.
- **Scenario 8:** Increasing dosing of devices per session: all patients receive the maximum 84mg (3 devices) at every visit. We have provided a highly conservative assumption that all patients would receive the maximum ESK-NS dosage (number of devices) at each administration session.
- **Scenario 9**: Using the ERG cap corrected for face validity (i.e., using 16.8% as relapse cap instead of original 12.8% cap), but capping relapse and loss of response to 1st line model inputs.
- **Scenario 10**: Adjusting both ESK-NS and OAD efficacy per the SUSTAIN-2 relative treatment effect across lines of treatment between the 3+ prior OAD subgroup and 3+ prior OAD + augmentation subgroup).
- **Scenario 10b**: Only adjusting OAD efficacy per the SUSTAIN-2 relative treatment effect (between the 3+ prior OAD subgroup and 3+ prior OAD + augmentation subgroup), and keeping ESK-NS efficacy to 3+ prior failures subgroup inputs (from TRANSFORM trials)
- **Scenario 11**: Keeping efficacy from 3+ prior failure subgroup (TRANSFORM-2/-3). This scenario uses the previous 3+ OAD failures efficacy, instead of the adjusted ESK-NS efficacy using the SUSTAIN-2 data.

Table 14: Population: 3+ OAD failures plus augmentation (Scenario 1-6)

Model Input	Scenario 1: No carer disutility	Scenario 2: Discontinuation in recovery 50%	Scenario 3: Shorter time horizon (5 years)	Scenario 4: TRD costing (75%) and Byford (25%) for health state costs	Scenario 5: nurse monitoring 2:1	Scenario 6: Excess mortality from Cuijpers source
Comparator	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC	No change from BC
Time Horizon	20 years	20 years	**5 years**	20 years	20 years	20 years
Discount rate	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)
Utility	No change: In line with 3rd ACM model	3 oAD + population utility	4th line + population utility	4th line + population utility	4th line + population utility	4th line + population utility
Carer Disutility	**Carer disutility excluded**	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included
Subsequent Treatment Efficacy & LT Outcomes	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap
Dosing Schedule	Increase in devices/session	Increase in devices/session	Increase in devices/session	Increase in devices/session	Increase in devices/session	Increase in devices/session
Administration Costs	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)	**2:1 (£50.92)**	1:1 (£72.92)**)**
Health State Costs	TRD cost study	TRD cost study	TRD cost study	**75% TRD/ 25% Byford**	TRD cost study	TRD cost study
Non-efficacy discontinuation of patients in recovery by 2 years	60%	**50%**	60%	60%	60%	60%
Excess Mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	**Cuijpers (1.52)**
ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 15: Population: 3+ OAD failures plus augmentation (Scenario 7-11)

Model Input	Scenario 7: Discount rate to 1.5%	Scenario 8: dosing max for devices and sessions	Scenario 9: ERG cap corrected for face validity	Scenario 10: Applying SUSTAIN-2 proportional efficacy reduction to both ESK-NS and OAD	Scenario 10b: Applying SUSTAIN-2 proportional efficacy reduction to OAD	Scenario 11: using 3+ TF inputs
Comparator	No change from BC	No change from BC	No change from BC	**OAD:** [redacted] **ESK-NS:** [redacted]	**OAD:** [redacted] **ESK-NS:** [redacted]	**ESK-NS and OAD efficacy informed from TRANSFORM 3+ prior input subgroup**
Time Horizon	20 years	20 years	20 years	20 years	20 years	20 years
Discount rate	**1.5% (for costs and health)**	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)	3.5% (for costs and health)
Utility	4th line + population utility	4th line + population utility	4th line + population utility	4th line + population utility	4th line + population utility	4th line + population utility
Carer Disutility	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included	Carer disutility included
Subsequent Treatment Efficacy & LT Outcomes	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	**ERG revised) but cap to 1st line model inputs (16.8% relapse, 23.1% loss of response)**	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap	ERG cap (error revised) to 1st line cap
Dosing Schedule	Increase in devices/session	**Maximum devices per administration***	Increase in devices/session	Increase in devices/session	Increase in devices/session	Increase in devices/session
Administration Costs	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)	1:1 (£72.92)
Health State Costs	TRD cost study	TRD cost study	TRD cost study	TRD cost study	TRD cost study	TRD cost study
Discontinuation of patients in recovery	60%	60%	60%	60%	60%	60%
Excess Mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality	No excess mortality
ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*For this scenario, the number of devices in Weeks 1-4 used an average of 2.875 devices, and all other weeks was 3.0 devices. This is reflective of the maximum average number of devices as per label wording (first dose should be 56mg, 2 devices).

Summary of results from the base case and scenario analyses

The base case ICERs indicate for people who have failed 3+ OAD failures and for people who have failed 3+ OAD failures and after augmentation, that ESK-NS is a cost-effective use of resources, with the ICERs being [redacted] per QALY and [redacted] per QALY respectively. Nevertheless, we conducted several univariate scenario analyses at their extremes in order to provide conservative ICER estimates to the committee, and to determine sensitivities around model parameters.

In the 3+ OAD failures scenarios overall, the results in the majority of scenarios we tested demonstrated that ESK-NS is cost effective, since the results were below the cost-effectiveness threshold that is accepted by NICE. The two parameters which had the biggest impact on the ICERs were the assumption that all patients receive the highest dosage of ESK-NS for all treatment sessions and cost inputs being sourced 50:50 from the Byford study, which is not appropriate since the costs are not based on a TRD population (see section 3). In the 3+ OAD failures and after augmentation population, the ICER results were [redacted] than those seen in the 3+ OAD failures population and could also be seen as a cost-effective use of resources given all scenarios were below the cost-effectiveness threshold that is accepted by NICE.

6. Summary and conclusions

In summary, Janssen has provided the committee with an updated base case analysis in two sub populations: both after 3+ OAD failures and after 3+ OAD failures and augmentation. The sub population 2 (after at 3+ OAD failures and after augmentation) position is aligned with feedback from stakeholders regarding the likely use of ESK-NS and where there is a significant unmet need for people with TRD. This positioning should substantially mitigate the concerns raised around implementation. New evidence provided is supportive of the original evidence presented in the submission demonstrating consistent efficacy and safety for ESK-NS. In addition, new evidence supports several key modelling assumptions such as the inappropriateness of Byford to inform the resource use for the health states and the capping approach used for the relapse rate for subsequent treatments in the model. This accompanied with a significantly revised PAS means that the ICERs are below NICE's accepted level of costs effectiveness and robust to the key scenarios, which show the cost effectiveness remains below £20,000 per QALY in all scenarios tested.

Janssen acknowledge that there are remaining uncertainties in the evidence base, and which is reflective of challenges on conducting mental health research generally and the lack of significant research in TRD as a condition, especially regarding the long-term outcomes. We ask the committee to bear this in mind in coming to their decision. We note the uncertainty in the evidence provided for sub population 2 but note that that this similar approach has been applied pragmatically in both the update NICE guideline for depression and TA 367 vortioxetine for treating major depressive episodes. Janssen believe that a recommendation in 3+ OAD failures, and especially in a 3+ OAD failures and augmentation population, offers an opportunity to make ESK-NS available to those people who need it most, while allowing a manageable implementation that allows for additional data to be collected that could lead to a wider recommendation upon a future NICE review. We hope that the committee are able to recommend ESK-NS to allow a new mechanism of action to be made available for people with TRD in England and Wales for the first time in 30 years.

References

[1] U. Government, “Health and Social Care Act 2012,” 2012. [Online]. Available: https://www.legislation.gov.uk/ukpga/2012/7/contents/enacted.

[2] C. H. Wong, K. W. Siah and A. W. Lo, “Estimation of clinical trial success rates and related parameters,” *Biostatistics,* vol. 20, no. 2, pp. 273-286, April 2019.

[3] A. Cipriani and J. Geddes, “Placebo for depression: we need to improve the quality of scientific information but also reject too simplistic approaches or ideological nihilism,” *BMC Medicine,* vol. 12, no. 105, pp. 1741-7015-12-105, 2014.

[4] B. Gaudiano and J. Herbert, “Methodological issues in clinical trials of antidepressant medications: perspectives from psychotherapy outcome research,” *Psychotherapy and Psychosomatics,* vol. 74(1), pp. 17-25, 2005.

[5] NICE, “Depression in adults: treatment and management consultation (update) CG90,” November 2021. [Online]. Available: https://www.nice.org.uk/guidance/gid-cgwave0725/documents/html-content-3.

[6] N. Kubitz, M. Mehra, R. Potluri, N. Garg and N. Cossrow, “Characterization of treatment resistant depression episodes in a cohort of patients from a US commercial claims database,” *PLoS One,* vol. 18;8(10), 2013.

[7] J. Reutfors, T. Andersson, P. Brenner, L. Brandt, A. DiBernardo, G. Li, D. Hagg, L. Wingard and R. Boden, “Mortality in treatment-resistant unipolar depression: A register-based cohort study in Sweden,” *Journal of Affective Disorders,* vol. 1;238, pp. 674-679, 2018.

[8] R. Feldman, D. Dunner, J. Muller and D. Stone, “Medicare patient experience with vagus nerve stimulation for treatment-resistant depression,” *Journal of Medical Economics,* pp. 6(1):62-74, 2012.

[9] K. Johnston, L. Powell, I. Anderson, S. Szabo and S. Cline, “The burden of treatment-resistant depression: A systematic review of the economic and quality of life literature,” *Journal of Affective Disorders,* pp. 242:195-210, 2019.

[10] D. Jaffe and T. Denee, “PMH73 - The Burden of Treatment-Resistant Depression in Europe from the Pationt Perspective,” in *ISPOR*, Barcelona, Spain, 10-14 November 2018.

[11] T. Denee, T. Ming, J. Waller, T. Bailey, O. Rajkovc-Hooley, C. Middleton-Dalby, H. Le, Q. Zhang, P. McCrone and D. Taylor, “A retrospective chart review study to quantify the monthly medical resource use and costs of treating patients with treatment resistant depression in the United Kingdom,” *Psychiatry,* pp. 311-319, 2020.

[12] J. Rush, M. Trivedi, S. Wisniewski, A. Nierenberg, J. Stewart, D. Warden, G. Niedereh, M. Thase and P. Lavori, “Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report,” *The American Journal of Psychiatry,* pp. 163(11):1905-17, 2006.

[13] D. A. Mrazek, J. C. Hornberger, A. C. Altar and I. Degtiar, “A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013,” *Psychiatric Services,* vol. 1;65(8), pp. 977-87, 2014.

[14] NICE, “Esketamine for treatment-resistant depression (ID1414): Appraisal consultation 2,” August 2020. [Online]. Available: https://www.nice.org.uk/guidance/gid-ta10371/documents/129-2.

[15] E. Daly, “A randomized withdrawal, double-blind, multicenter study of esketamine nasal spray plus an oral antidepressant for relapse prevention in treatment-resistant depression.,” in *The Annual Meeting of the American Society of Clinical Psychopharmacology*, Miami, Florida, 2018.

[16] V. Popova, E. Daly, M. Trivedi, K. Cooper, R. Lane, P. Lim, C. Mazzucco, D. Hough, M. Thase, R. Shelton, P. Moleo and E. Vieta, “Efficacy and Safety of Flexibly Dosed

Esketamine Nasal Spray Combined With a Newly Initiated Oral Antidepressant in Treatment-Resistant Depression: A Randomized Double-Blind Active-Controlled Study," *The American Journal of Psychiatry,* vol. 176(6), pp. 428-438., 2019.

[17] D. Jaffe, B. Rive and T. Denee, "The humanistic and economic burden of treatment-resistant depression in Europe: a cross-sectional study," *BMC Psychiatry,* vol. 9(1):247, 2019.

[18] NICE, "Esketamine for treatment-resistant depression (ID1414): Committee papers SANE response," 28 January 2020. [Online]. Available: https://www.nice.org.uk/guidance/gid-ta10371/documents/committee-papers.

[19] NICE, "Vortioxetine for treating major depressive episodes: TA367," 25 November 2015. [Online]. Available: https://www.nice.org.uk/guidance/ta367.

[20] T. Denee, C. Kerr, T. Ming, R. Wood, T. Tritton, C. Middleton-Dalby, O. Massey and M. Desai, "Current treatments used in clinical practice for major depressive disorder and treatment resistant depression in England: A retrospective database study," *Journal of Psychiatric Research,* vol. 139, pp. 172-178, 2021.

[21] M. Dold, L. Bartova and S. Kasper, "Treatment Response of Add-On Esketamine Nasal Spray in Resistant Major Depression in Relation to Add-On Second-Generation Antipsychotic Treatment," *International Journal of Neuropsychopharmacology,* vol. 23, no. 7, pp. 440-445, 2020.

[22] B. Wright, E. Eiland 3rd and R. Lorenz, "Augmentation with atypical antipsychotics for depression: a review of evidence-based support from the medical literature," *Pharmacotherapy,* pp. 344-59, 2013.

[23] NICE, "Esketamine for treatment-resistant depression (ID1414): Janssen Responses to ACD 2 (Table 15)," 2020. [Online]. Available: https://www.nice.org.uk/guidance/gid-ta10371/documents/committee-papers-2.

[24] Janssen, "Esketamine for treatment-resistant depression (ID1414): ACD response - Janssen Readiness Survey," 2020. [Online]. Available: https://www.nice.org.uk/guidance/gid-ta10371/documents/html-content-3.

[25] J. r. t. A. 2, "Esketamine for treatment-resistant depression (ID1414): ACD Consultation 2," 2020. [Online]. Available: https://www.nice.org.uk/guidance/indevelopment/gid-ta10371/documents.

[26] E. Waja, L. Aluisio, R. Holder, E. Daly, R. Lane, P. Lim, J. George, R. Morrison, G. Sanacora, A. Young, S. Kasper, A. Sulaiman, C.-T. Li, J.-W. Paik, H. Manji, D. Hough, J. Grunfeld, H. J. Jeon, S. Wilkinson, W. Drevets and J. Singh, "Esketamine Nasal Spray Plus Oral Antidepressant in Patients with Treatment Resistant Depression: Assessment of Long-Term Safety in a Phase 3, Open Label Study (SUSTAIN-2)," *Journal of Clinical Psychiatry,* p. 81/3;, 2020.

[27] Janssen, "Safety and Sustenance of Esketamine Treatment Response with Repeated Doses at Intervals Determined by Symptom Severity (SUSTAIN-3): Interim CSR," December 2021.

[28] A. Fuertes-Saiz, V. Elvira, J. Etxeandia-Pradera, M. Vega and R. Mora-Marin, "Effectiveness and Tolerability of Esketamine Nasal Spray in Treatment-Resistant Depression: A Spanish Nationwide Study of its Compassionate Use," in *Americal Psychiatric Association*, 2021.

[29] L. Samalin, L. Mekoui, M.-A. Codet, J. Boursicot-Beuzelin and E. Gaudre-Wattinne, "French real-world study (ESKALE) of esketamine for patientswith resistant depression: CO4-2," in *L'ENCEPHALE*, Paris, France, January 2022.

[30] S. Byford, B. Barrett, N. Despiegel and A. Wade, "Impact of treatment success on health service use and cost in depression: longitudinal database analysis," *Pharmacoeconomics,* vol. 29(2), pp. 157-170, 2011.

[31] S. Bradley, N. Lawrence and P. Carder, “Using primary care data for health research in England – an overview,” *Future Healthcare Journal,* vol. 5(3), p. 207–212, 2018.

[32] T. Kendrick, J. Chatwin, C. Dowrick, A. Tylee, R. Morriss, R. Peveler, M. Leese, P. McCrone, T. Harris, M. Moore, R. Byng, G. Brown, S. Barthel, H. Mander, A. Ring, V. Kelly, V. Wallace, M. Gabbay, T. Craig and A. Mann, “Randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of selective serotonin reuptake inhibitors plus supportive care, versus supportive care alone, for mild to moderate depression with somatic symptoms in primary care,” *Health Technology Assessment,* vol. 13(22), pp. 1-159, 2009.

[33] P. Fonagy, F. Rost, J.-A. Carlyle, S. McPherson, R. Thomas, R. P. Fearton, D. Goldberg and D. Taylor, “Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock Adult Depression Study (TADS),” *World Psychiatry,* vol. 14, no. 3, pp. 312-21, 2015.

[34] H. Sacheim, “Acute Continuation and Maintenance Treatment of Major Depressive Episodes With Transcranial Magnetic Stimulation,” *Brain Stimulation,* vol. 9, no. 3, pp. 313-319, 2016.

[35] B. Wu, Q. Ca, J. Sheehan, C. Benson, N. Connolly and L. Alphs, “An episode level evaluation of the treatment journey of patients with major depressive disorder and treatment-resistant depression,” *PLoS One,* vol. 8;14(8), 2019.

[36] K. Heerlein, G. Perugi, C. Otte, T. Frodl, G. Degraeve, W. Hagedoorn, A. Oliveira-Maia, V. Perez Sola, S. Rathod, G. Rosso, P. Sierra, S. Malynn, J. Morrens, C. Vrrijcken, B. Gonzalez and A. Young, “Real-world evidence from a European cohort study of patients with treatment resistant depression: Treatment patterns and clinical outcomes,” *Journa of Affective Disorders,* vol. 1, no. 290, pp. 334-334, 2021.

[37] J. Talbot, D. MacIntyre, S. Rathod, J. Morrens and A. Young, “Treatment Resistant Depression in the UK: Sub-Analysis of a European Real-World Evidence Study,” in *Royal College of Psychiatrists International Congress*, 2021.

Esketamine for treatment-resistant depression

ERG response to ACD3

Produced by Kleijnen Systematic Reviews Ltd (KSR)

Authors
Robert Wolff, Managing Director, KSR
Nigel Armstrong, Health Economics Manager, KSR
Caro Noake, Information Specialist, KSR
Jos Kleijnen, Founder and Owner, KSR

Correspondence to
Robert Wolff, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 14/03/2020

Source of funding:
This report was commissioned by the NIHR HTA Programme as project number 12/78/96.

Declared competing interests of the authors
None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Rider on responsibility for report

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

Contributions of authors

Robert Wolff acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Nigel Armstrong acted as health economic project lead, critiqued the company's economic evaluation, and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report, and supervised the project.

Abbreviations

ACD	Appraisal Consultation Document
AiC	Academic in confidence
BSC	Best supportive care
CBT	Cognitive behavioural therapy
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CiC	Commercial in confidence
ECT	Electroconvulsive therapy
EMA	European Medicines Agency
EQ-5D-5L	European Quality of Life-5 Dimensions – 5 levels
ERG	Evidence Review Group
ESK	Esketamine
ESK-NS	Esketamine nasal spray
FDA	U.S. Food and Drug Administration
HCRU	Healthcare resource use
HDRS	Hamilton Depression Rating Scale
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
KSR	Kleijnen Systematic Reviews
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	Major depressive disorder
MDE	Major depressive episode
MHRA	Medicines and Healthcare products Regulatory Agency
MHT	Mental health trust
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NS	Nasal spray
OAD	Oral antidepressant
PAS	Patient access scheme
PHQ-9	Patient Health Questionnaire – 9 questions
QALY	Quality-adjusted life year
QIDS	Quick Inventory of Depressive Symptomatology
QoL	Quality of life
TRD	Treatment-resistant depression
UK	United Kingdom

Table of Contents

1. **Submitted new evidence to support ESK-NS in a new optimised population who have failed 3 or more OADs and after augmentation addresses a high unmet need and is supported by clinical opinion and available evidence (which demonstrates that ESK-NS would be a clinically and cost-effective option)**

Based on the FAD, and in particular clinical expert feedback, the company have submitted evidence for two populations (p.7):

- After failure of 3 or more oral antidepressants (as presented at the 3rd ACM) – *"3+ OAD failures"*
- After the failure of 3 or more oral antidepressants and after 1 previous augmentation with an atypical antipsychotic or lithium – *"3+ OAD failures and after augmentation"*

1.1 The use of ESK-NS in the 3+ OAD failures and after augmentation subgroup fulfils a significant unmet need and is an appropriate position in the pathway for the committee to consider the clinical and cost effectiveness of ESK-NS

The company provide evidence for [redacted] from the [redacted] **ERG comment:** The ERG notes the general trend of increasing resource use and duration of depression and disease burden. However, the units in Table 2 seem to be misreported as months instead of numbers per patient per unit time. Also, the differences by line of TRD are generally not that large and there are some notable exceptions to the trend with the number of non-elective and elective admissions being very slightly lower for at least 4 lines of therapy than for at least 2 lines: [redacted] and [redacted] vs. [redacted] and [redacted] This seems to resonate with the lack of increase that appears to be between later lines of therapy observed in the cost study by Denee et al. 2021[1] (see 3.2 below).

1.2 To inform this new revised positioning, evidence of the efficacy of ESK-NS+OAD in the 3+ OAD failures subgroup has been adjusted to a population that have failed 3+ OADs and augmentation subgroup using available evidence and an approach previously accepted by NICE

Estimates of remission and response for each of these populations was presented in Table 4: those for 3+ OAD failures, estimated from the TRANSFORM studies, were originally presented in response to ACD and reported in Table 2 of the ERG critique. Those for ESK-NS + OAD in 3+ OAD failures and after augmentation were stated to have been calculated by applying the *"relative treatment effect between SUSTAIN-2 3+ prior OAD and 3+ prior OAD and augmentation subgroups"* to the TRANSFORM values (Table 4).

ERG comment: It is unclear to the ERG why the company chose to use the relative treatment effect from SUSTAIN-2 instead of the estimates of the 3+ prior OAD and augmentation subgroup from the TRANSFORM studies. It might be that those data were not available, or the sample size deemed to be

too small, but the ERG can find no explanation. The ERG can confirm that the relative treatment effect was the difference in values between the two subgroups as a percentage of the 3+ OAD failures subgroup, as reported in Table 3, Appendix C.[2] However, applying these percentages to the TRANSFORM values, the ERG could not reproduce the results in Table 4 precisely: [redacted] instead of [redacted] for remission and [redacted] instead of [redacted] for response. However, the ERG considers that the approach taken is probably conservative given that no adjustment was made to the values for OAD, despite some reduction between 2+ OAD failures and 3+ OAD failures lines in TRANSFORM-2 and between Step 3 and Step 4 in the STAR*D trial. How conservative is difficult to be sure given that it is unclear what the effect of augmentation might be from the evidence provided: as stated above, ideally data for this subgroup would have come from the TRANSFORM RCTs.

1.3 Monotherapy OAD remains the main comparator for ESK-NS in a 3+ OAD and after augmentation population with treatment effect for an OAD conservatively assumed as maintained between lines of treatment

The company argues that OAD should be the comparator because of the lack of evidence on other comparators and because of limited use of augmentation. They state that *"meta-analyses have shown that ESK-NS had a relative effect size that was nearly twice as high versus antidepressant augmentation with second-generation antipsychotics than compared with a monotherapy OAD"* and *"the adverse effect profile of augmentation treatments may be unfavourable to some patients"* (p.12). They also argue that assuming OAD is the comparator is conservative given that no additional cost of augmentation is incurred for the comparator.

ERG comment: It is unclear why augmentation is not the appropriate comparator for 3+ OAD failures given that it forms the basis of the other population, suggesting that it would be the next line of therapy. It is also not clear from the evidence presented what the next line of therapy post-augmentation might be and therefore unclear that it would be OAD and therefore that OAD would also be the comparator for the post-augmentation population. Of course, if OAD is the appropriate comparator then it is likely that the approach to not adjust the remission and response rates for the later line subgroup would appear to be conservative. However, if the appropriate comparator is not OAD then the effectiveness of the comparator in both subgroups could be underestimated. Indeed, the company claim that the treatment effect vs. OAD was greater for ESK-NS than for augmentation with second-generation antipsychotics This is based on the study by Dold et al. 2020, which showed a MSDRS treatment effect vs. placebo of pooled add-on esketamine nasal spray trials (n=3, n=641; MD=4.09, 95% CI: 2.01 to 6.17) vs. the pooled SGA augmentation trials (n=23, n=8363; MD=2.05, 95% CI: 1.51 to 2.59).[3] However, this simply highlights that augmentation is more effective than OAD monotherapy and thus the treatment effect of ESK-NS versus augmentation would be lower than versus OAD monotherapy.

1.4 The 3+ OAD failures subgroup efficacy is supported by the evidence from open label ESK-NS clinical trials and emerging RWE evidence

The company presented evidence from SUSTAIN-2 to validate the remission and response rates from TRANSFORM-2 in the 3+ OAD failures population, as stated to have been reported in Table 5.

ERG comment: The values from SUSTAIN-2 could be regarded as similar to shoes from TANSFORM-2: [redacted] vs. [redacted] and [redacted] vs. [redacted] for remission and response respectively. Table is labelled as relating to *"3+ OAD failures and after augmentation subgroups"*, but the values for SUSTAIN-2 do correspond with those in Table 2 of Appendix C for 3+ OAD failures and the ERG can also confirm that the figures for

TRANSFORM-2 are the ones originally presented by the company in Table 8 of their response to ACD 2.

1.5 The 3+ OAD failures and after augmentation population significantly improves previous implementation concerns given the smaller eligible population and due to patients already likely to be receiving care in a secondary mental health setting

The company argued that because patients at this line of therapy are already receiving care in the secondary care setting there would be little extra cost associated with the introduction of ESK-NS: *"The new positioning of ESK-NS in the pathway after augmentation means that virtually all patients will be managed within a secondary care setting given that augmentation is not a primary care intervention."* (p.15) The company also estimated the number of eligible patients at this line would be 14,745 based on Denee at al.[1]

Nevertheless, they decreased the patient: nurse ratio in the model for this population from 2:1 to 1:1 (See 5.1 below).

ERG comment: Any lack of increase in cost associated with ESK-NS, as the company point out, is predicated on that cost already being incurred. However, the main concern of the committee regarding additional cost, as mentioned by the company, was the need for changes in infrastructure. It is unclear to the ERG how patients having been prescribed augmentation in secondary care would affect the need for change in infrastructure due to the introduction of ESK-NS.

2. New evidence from the long term ESK-NS safety studies shows that ESK-NS has a manageable safety profile, while emerging RWE demonstrates a consistent efficacy profile to the esketamine RCTs.

In response to [redacted] the company presented new evidence, namely

- The phase 3 clinical trials, SUSTAIN-2 and SUSTAIN-3 (see Section 2.1).
- Two RWE studies from a French and Spanish cohort receiving ESK-NS in clinical practice (see Section 2.2).

2.1 Safety Data Update (SUSTAIN-2 and SUSTAIN-3): demonstrates that esketamine has a favourable safety profile with acceptable tolerability

The company presented the findings of two studies, SUSTAIN-2 and SUSTAIN-3, which are both open-label.

2.1.1 SUSTAIN-2

Out of the 802 patients enrolled in SUSTAIN-2, 691 were direct-entry patients, and 111 were transferred from the 3005 study (ESKETINTRD3005, TRANSFORM-3). According to the company, "*long-term safety of intermittent treatment with esketamine plus an oral antidepressant was favourable with acceptable tolerability. The most clinically relevant safety findings were transient (i.e., resolving on the day of esketamine administration or within 1.5 hr of administration)*".

The company stated that "*most adverse events were tolerated, being either mild or moderate in intensity. Of the patients ([redacted]) with severe adverse events during treatment, the most common (≥1%) included those related to the disease itself (i.e., TRD) (anxiety) or were related to esketamine administration (dissociation, dizziness, nausea, and anxiety). The majority of severe adverse events occurred on the day of esketamine administration were transient (i.e., 75% of these events resolved on the same day). Serious adverse events were reported in 6.9% of patients (55 of 802) during treatment in this 1-year study. The most common serious adverse events (SAEs) (in ≥2 people) were depression (1.0%), suicidal ideation (0.7%), suicide attempt (0.7%), anxiety (0.2%), and gastroenteritis (0.2%)"*.

Furthermore, *"there were 2 deaths during the study, both occurring during the optimisation/maintenance phase. One patient died as a result of a completed suicide, and the other patient as a result of cardiac/respiratory failure. The cases were considered not related or doubtfully related to esketamine treatment, respectively, according to the investigator and not related to esketamine treatment by the sponsor"*.

The company further highlighted that *"the events reported after discontinuation of treatment from the induction phase may have reflected a return of depressive symptoms in patients who had been experiencing a partial response to esketamine treatment:* [redacted]*"*.

ERG comment: SUSTAIN-2 included participants of TRANSFORM-3, a study that included adults aged 65 years and over. While SUSTAIN-2 appeared to be a well conducted observational study, it is

a non-comparative open-label study and as such will be open to bias. Both, SUSTAIN-2 and TRANSFORM-3 were not initially included in the economic model (see Section 4.2.1 of the ERG report).

As highlighted in the ERG report, *"the ERG is concerned with the lack of clarity on dosing in TRANSFORM-2 and TRANSFORM-3 trials plus the complex dose changes in SUSTAIN-1 and SUSTAIN-2"*. It should also be noted that for SUSTAIN-2 [redacted]. No details have been given regarding the nature of the issue.

Overall, the ERG remains to be concerned regarding the safety data of esketamine.

2.1.2 SUSTAIN-3

According to the company, *"SUSTAIN-3 is a global multi-centre trial, across 222 sites and in 27 countries, including 4 sites in the UK. Adult patients (≥18 years) who previously participated in 1 of 6 of the phase 3 ESK-NS trials, or "parent" studies of ESK-NS are enrolled. Each parent study enrolled patients with either recurrent or a single-episode (≥2 years) of major depressive disorder without psychotic features and who met the definition of TRD"*.

Regarding serious adverse events, the company states that these [redacted].

According the company, [redacted].

ERG comment: [redacted].

In line with comment 3.14, uncertainty about long-term saftey might have been partially resolved by including results of the SUSTAIN-3 trial. However, the concerns raised in the comment, especially regarding [redacted] are not fully resolved.

2.2 The French ESKALE study and Spanish compassionate use program provide supportive evidence in a real-world population regarding overall treatment efficacy for ESK-NS

In response to Section 3.12 of the FAD (The evidence from the trials is limited in its generalisability to the NHS), the company stated that *"since the last ACM, initial results of a French real-world evidence study (ESKALE) of ESK-NS for patients with TRD and initial results from a Spanish compassionate use programme of ESK-NS have become available. These data, despite initial use in populations likely to be significantly more severe and treatment resistant than populations studied in the ESK-NS clinical trial program given they were mainly used in compassionate use programs, nevertheless, provide supporting information on the efficacy and safety of ESK-NS"*.

2.2.1 ESKALE study

ESKALE included 160 patients from France and was a retrospective, observational study of adults aged 18 years and over, with moderate-to-severe TRD, (where TRD is defined as non-response to ≥2 oral antidepressant drugs). Results of an interim anaylsis after six months of follow-up are presented.

"After 6 months of treatment with ESK-NS, patients experienced a decrease of 16.0 points in median MADRS total score from baseline, which corresponds to 46.7% of responders and 26.7% of remitters".

However, as highlighted by the company, e.g. due to low number of observations (n=30) and patients (n=14), *"current results at 6 months should be interpreted with caution until further data is collected"*.

ERG comment: It should be noted that [redacted] of the FAD noted that [redacted] and [redacted].

While limited information has been provided in Table 6 in Appendix F of the company response, e.g. regarding *"lifetime suicide attempts"*, it is unclear how the provided results can overcome the concerns regarding the generalisability to NHS patients in England and Wales as well as to address the concern regarding the exclusion of a [redacted].

2.2.2 Spanish ESK-NS compassionate use program

According to the company, *"a total of 32 patients were included, with a mean age of 54.9 years and 69.9% who were females. ECT had been used in 46.9% (n=15) and of the remaining 53.1% (n=17)*

patients, 58.8% (n=10) rejected the treatment, 29.4% (n=5) did not have access to this procedure and in 11.8% (n=2) of cases ECT was partially contraindicated. This cohort was a significantly sicker cohort aligned with compassionate use. ESK-NS was effective in 87.5% (n=28) of patients with response and remission rates after 6 months being 56.3% (n=18) and 31.3% (n=10), respectively. The majority of responders, 55.6% (n=10) responded during the first week and 22.2% (n=3) during the first month. Adverse events were mild, and tolerability was good with dizziness in 15.6% of patients (n=5), dissociative symptoms in 9.4% (n=3), anxiety in 3.1% (n=1) and 71.1% (n=23) reported no adverse effects)".

ERG comment: For these data, the concerns of the ERG are as outlined in Section 2.2.1.

3. A new analysis of UK data demonstrates the high resource utilisation and high healthcare cost of treating people with TRD, supporting the original TRD cost study rather than the Byford et al study, which is not appropriate given the setting of care.

3.1 Using the Byford et al study solely significantly underestimates the NHS cost of treating patients with TRD

The company argue that the Byford et al. study[4] is not appropriate to estimate the health state costs in the model for two main reasons: primary care population and not TRD.

ERG comment: The ERG would argue that just because patients were identified using a primary care source (GPRD database) does not mean that secondary care costs were not included and that costs for TRD are not all incurred in the secondary care setting. However, it is true that the nearest match to TRD was 'severe depression' and that this appears to be a heterogenous group, including diagnoses of unclear resource implication such as 'Endogenous depression – recurrent' as well as psychosis, which is likely to be quite resource intensive.[4] It is also the case that the cost of a MDE might increase consistent with the later lines of the latest evidence submission, although later line might only affect duration in the MDE state rather than cost per unit time in the state.

3.2 The significant NHS healthcare resources that patients with TRD utilise is highlighted in the TRD cost study and is further confirmed in a recently conducted retrospective database study

The company provided an estimate of the MDE health state cost based on the cost study, by Denee et al. 2021,[1] based on the Clinical Record Interactive System (CRIS) produced by the South London and Maudsley NHS Foundation Trust (SLaM). According to Appendix G, [redacted].[2]

[redacted] This increased to [redacted] respectively.

ERG comment: The TRD study does seem to provide an estimate of cost according to definitions of TRD that have some merit i.e. according to line of therapy, which is more appropriate than using some

notion of severity as with the Byford et al study. However, the ERG has several concerns regarding the accuracy of the TRD estimates.

- Firstly, as indicated by the figures for precision and sensitivity, TRD might be misidentified. Lack of precision is probably more serious given that TRD patients will be contaminated by non-TRD patients, and the ERG have located some notes provided by the SLaM that suggest that it might be much lower: *"67% is a more representative figure of the app's precision performance"* (Section 6. TREATMENT-RESISTANT DEPRESSION, CRIS NLP SERVICE Library of production-ready applications UPDATED 15/10/2021).[5]
- Secondly, the median and interquartile range limits for inpatient bed nights [redacted] despite this cost being of such a high proportion of total cost e.g. [redacted] out of [redacted] for 6 months.[2] This suggests that inpatient bed days is very skewed and so might be particularly prone to bias, including by including non-TRD patients.
- Thirdly, if treatment resistance is considered to be the main driver of cost, as opposed to any notion of severity or any other characteristics of the depression e.g. suicidal ideation or psychosis, then it might be expected that cost will increase with line of therapy. However, although cost increases in moving from TRD definition 1 to 2, there is no increase in moving from definition 2 to 3.
- Fourthly, there is no information on other patient characteristics that are likely to be cost drivers such as psychosis or suicidal ideation.[2] The combination of these concerns is that the sample of patients from which the latest MDE cost estimates are derived might include patients who are not typical of TRD and might produce a bias in the estimates.

4. The ERG treatment cap in the model and subsequent treatment efficacy is overestimating long-term outcomes for people with TRD based on the literature

4.1 The current ERG treatment cap based on Wu et al has been corrected to increase face validity, but remains a highly conservative estimate of subsequent treatment efficacy

The company argued that the study, by Wu et al.,[6] used by the ERG to validate time spent in the MDE state, is less relevant because it is based in the US and because: "[redacted]"(p. 27) They reported that the study showed that the mean length of the first TRD episode was 1.56 years, and the mean length of remission was 0.90 years and that "*The clinical expert noted at the 3rd ACM that this is likely to be optimistic and commented that he felt the true estimate may lie between the ERG and Janssen's assumptions."* (p. 28)

4.2 In additional to previous evidence submitted ahead of the 3rd ACM, new evidence suggests that a proportion of TRD patients are likely to spend a significant period of time in MDE due to low levels of remission and high levels of relapse

The company also present additional evidence in the form of a study based on the DISCOVER dataset, based on UK patients, to provide alternative estimates as comparison to the 1.56 years in the Wu et al. study.[6] The results of the DISCOVER study for average depression duration for TRD (failed at least 2 lines of therapy) are: [redacted], and for patients who have failed 3 or more lines and 4 or more lines of treatment, [redacted], respectively. Also from the DISCOVER dataset, percentage remission and relapse for TRD were reported to be [redacted] respectively.

A median duration of 5 years (95% CI 4-6 years) and mean of 6.1 years are reported to be the duration of a TRD episode respectively in a UK study of 178 patients from 3 UK centres and of the 49 UK patients in a European study.[7]

ERG comment (on 4.1 and 4.2): The ERG recognises the limitation of the Wu et al. study that it was conducted in the US.[6] It is also true that follow-up was limited to include no more than two episodes and it is unclear that proportion of second episodes were of TRD in those whose first episode was of TRD. However, it is unclear to the ERG why the results from this study of duration of episodes should not be [redacted]. The new evidence presented by the company does suggest a longer duration of depressive episode, which does support the clinical expert opinion that the Wu et al. estimate was optimistic, particularly given that the new evidence is from UK patients. However, the difference is so large that the ERG wonders whether the Wu et al. and company cited studies are measuring the same thing and whether there is a conflation between definitions of episode, one of which might include periods of time not suffering from depression and the other that is a continuous period in a depressed state, Indeed, the company provided values seem to bear a greater resemblance to total time spent in the MDE health state with periods in and out of the MDE health state between lines of therapy. The ERG calculation of proportion of time spent in the MDE state is based on a 20-year time horizon, which implies a life expectancy of about 13.8 years with time in the MDE state of about 6.6 years according to the ACD 2 preferences, albeit with implausible value for loss of response and 7.1 years with the ERG cap applied.[8] As shown by the ERG description of the calculation of proportion of time in the MDE state over the 20 year time horizon, the time spent in the MDE state implicitly implies multiple 'episodes', but in the sense that there would be periods in and out of the MDE state. This is of course based on a heterogeneous population with patients taking many different

journeys. Some might response or remit first line TRD and never enter the MDE state again whereas, at the other extreme, some might never enter the remission state and remain in the MDE state. However, in the middle some would respond or remit and relapse several times with a change in treatment each time over that period. Those latter patients might still be regarded as being in the same episode of TRD in the sense that, on loss of response or relapse, they move to the next line of therapy, rather than restarting with first line treatment. Nevertheless, they will have had a period out of the MDE health state between lines of therapy. It seems that the Wu et al. study was provides estimates of time in state that are more consistent with 'episode' as defined by time continuously spent in a particular health state, either MDE or remission, whilst the ERG hypothesises that the studies provided by the company provide estimates of time since diagnosis or first treatment of TRD including several periods in an MDE state with periods not in the MDE state in between. Some support for this also comes from the nature of the time estimate from the UK patients of the European study, which is at baseline and is accompanied by a report of 14.3% with at least five treatment failures over this period, which implies that many patients experience successful treatment and thus 'episodes' not in what might be regarded as a MDE state.[7] No further details of the DISCOVER data are provided to test this hypothesis.

4.3 The ERG cap has been corrected to maintain face validity in the model, and the amended ERG cap should be considered the upper limit of the cost effectiveness estimate given the literature

The company point out that the ERG cap implemented as part of the critique to the ACD 2 response resulted in values for relapse and loss of response of 22.8% and 12.8% for TRD lines 3 to 6 (BSC) that were lower than those at TRD line 2+, which appears to be implausible. The company therefore provide alternative estimates of the ERG cap for relapse and loss of response for TRD lines 3 to 6 (BSC for 3+ OAD) for both sub-populations i.e. 3+ OAD failures and at least 3+ OAD failures and after augmentation, which were 31.8% and 23.7% respectively. The method of calculation was reported in Appendix H,[2] which shows relapse as an example: 16.8%*(12.79%/6.77%) = 31.79%.

ERG comment: The original cap introduced by the ERG was arbitrary and the ERG agrees with the company that it was probably too low. The method chosen by the company is consistent with choosing the value of relapse or loss of response estimated for TRD line 2 by the ERG and according to the ACD 2 preferences.[8] Therefore, although still arbitrary, these values do appear to have greater validity. However, when used they results in a time in the MDE health state over the 20-year time horizon of 9.2 years, which is about 66% of life expectancy, which is 13.8 years. As the ERG argue above, it might be reasonable to spend this amount of time with a diagnosis of TRD if moving in and out of the MDE state, but this is not the same as being in the MDE state with no remission during this time. On this basis and notwithstanding the limitations of the Wu et al. study the ERG have produced a scenario with a new cap at TRD line 2+ (3+ failures) i.e. 16.82% for relapse and 23.05% for loss of response. It is important to note that the ERG is not arguing that these are the most plausible estimates per se, but instead, given the lack of evidence and questionable model validity, these are the values that produce a time in the MDE state that appears to be plausible given the Wu et al. study. This scenario was also one implemented by the company as Scenario 9 for 3+ OAD failures plus augmentation. This scenario is referred to as the ERG new cap (see 5.1 below).

5. A new base case is included for the committee's consideration: overall, the new evidence and updated model shows that ESK-NS is a cost-effective option in both the 3+ OAD failures, and the 3+ OAD failures and after augmentation position in the MDD pathway. The incremental cost effectiveness ratios (ICERs) with the new value proposition are below £10,000 per QALY threshold

5.1 Revised base case economic model inputs

The company provides two sets of analyses, including a base case and model, for each of the two new populations:

- **After failure of 3 or more oral antidepressants (as presented at the 3rd ACM) – "3+ OAD failures"**

The company stated that the only changes made to the model post-ACM 3 were:

1) Correction to ERG cap (see 4.3 above)
2) Health state costs are a weighted average of Byford et al (25%) and TRD cost study (75%) (see 3.2 above)

- **After the failure of 3 or more oral antidepressants and after 1 previous augmentation with an atypical antipsychotic or lithium – "3+ OAD failures and after augmentation"**

The company stated that the changes made to the model post-ACM 3 were:

1) Correction to ERG cap (see 4.3 above)
2) Health state costs solely from TRD cost study (100%) (see 3.2 above)
3) Short-term efficacy from TRANSFORM-2 adjusted for a population that have failed 3+ OAD failures and augmentation sourced from the SUSTAIN-2 trial induction period (see 1.2 above)
4) Dosing schedule increased to reflect the later line population where devices per session is increased based on extrapolation (see below)
5) Administration costs for nurse monitoring per patient revised from 2:1 to 1:1 to reflect reduced numbers of patients with the smaller population (See 1.5 above)
6) Percentage of patients in recovery who discontinue by 2 years due to non-efficacy reasons is decreased (from the 3+ OAD failures population) from 70% to 60% (see Appendix I)

The results of the base case analyses for these two populations alongside the results for the 3rd ACM 3+ OAD failures subgroup, but with the new PAS discount of [redacted] instead of [redacted] were reported in Table 12.

A revised PAS of [redacted] was submitted after completion of the critique of the company response to ACD 3.

ERG comment: The ERG were able to reproduce (within £2) the results of the 3rd ACM 3+ OAD failures subgroup, but with the new PAS discount of [redacted].

The ERG could also reproduce the results for 3+ OAD failures based on those two changes (revised ERG cap and weighted average health state cost): [redacted] in Table 12. With the revised PAS, the ICER is [redacted]. The company also implemented various scenarios, the one with the highest ICER, [redacted],

being that which assumed the maximum devices per session, although setting the number of devices to three, the ERG obtained an ICER of [redacted] or [redacted] with the revised PAS.
The ERG conducted a scenario analysis using 100% Byford costs, which produces and ICER of [redacted] or [redacted] with the revised PAS. Applying the new ERG cap increased the ICER to [redacted] or [redacted] with the revised PAS.

The ERG could also reproduce the results for 3+ OAD failures and after augmentation based on those six changes (revised ERG cap and weighted average health state cost): [redacted] in Table 12. With the revised PAS, [redacted]. The company also implemented various scenarios, the one with the highest ICER, [redacted], being referred to as ERG cap (error revised) to 1st line cap, which assumed the same cap as in the ERG scenario referred to as new ERG cap. The ICER for this scenario with the revised PAS is [redacted]. The ERG conducted a scenario analysis using 100% Byford costs, which produces and ICER of [redacted] or [redacted] with the revised PAS. Applying the new ERG cap increased the ICER to [redacted] or [redacted] with the revised PAS.

6. Summary and conclusions

The company have presented analyses for two populations:

- After failure of 3 or more oral antidepressants (as presented at the 3rd ACM) – *"3+ OAD failures"*
- After the failure of 3 or more oral antidepressants and after 1 previous augmentation with an atypical antipsychotic or lithium – *"3+ OAD failures and after augmentation"*

They have also argued that the comparator i.e. OAD remains the same, with no reduction in efficacy at the first line for each population with a reduction in efficacy for ESK-NS in moving to the 3+ OAD failures and after augmentation population. There has also been an increase to the relapse and loss of response rates to overcome the lack of plausibility of the so-called ERG cap. In addition, the company have provided some change from committee preferences to health state costs, either averaging the Byford et al.[4] and original CS TRD cost study estimates for the 3+ OAD failure population of just using the TRD cost study estimates for the 3+ failures and after augmentation population. Costs hae been adjusted for inflation and a new PAS discount has been incorporated. For the latter population the patient : nurse ratio had also been reduced from 2:1 to 1:1.

The ERG have been able to reproduce the new estimates of efficacy and the base case ICERs based on the changed reported by the company. The ERG would argue that there still remain issues regarding the validity of the model and its parameters for representing the natural course of the disease, in particular the time spend in the MDE state. The company have argued, based on some UK based data that the time in MDE state might be too optimistic. However, the ERG would tentatively suggest that the discrepancy between those UK based data versus the US based data presented by the ERG is so large as to suggest as difference in the definition of an episode of MDE: time since diagnosis/first treatment with periods in and out of and MDE state versus time continuously in and MDE state between lines of therapy respectively. This might mean that the US based data as applied by the ERG still retains some value in testing the validity of the model.

There also remains much uncertainty as to the size of health state costs. The new UK data provided by the company, based on Denee et al. 2021,[1] do suggest that cost is more in line with the original TRD. However, the ERG have some concerns regarding the definition of TRD in this UK study and the possibility of contamination with non-TRD patients or overrepresentation with patients with high resource consumption, potentially leading to selection bias. The data from another UK source, the [redacted], also cause doubt as the size of the increase in cost with line of TRD.

Finally, the ERG would also question the validity of the comparator given that the line of therapy following 3+ OAD failures is augmentation with an atypical antipsychotic or lithium. There is also some evidence to suggest that augmentation with an atypical antipsychotic or lithium might be more effective than OAD. Therefore, the treatment effect of ESK-NS for the two populations might be overestimated.

Notwithstanding the limitations of model structure and efficacy and cost parametrisation, the ERG have presented an additional two scenarios, one with original Byford et al. costs and one, in order to address the uncertainty in how the episodic nature of the condition is modelled, with an adjustment to the ERG cap on subsequent therapy loss of response and relapse.

References

[1] Denee T, Ming T, Waller J, Bailey T, Rajkovic-Hooley O, Middleton-Dalby C, et al. A retrospective chart review study to quantify the monthly medical resource use and costs of treating patients with treatment resistant depression in the United Kingdom. *Curr Med Res Opin* 2021;37(2):311-9.

[2] Janssen. *Esketamine for treatment-resistant depression [ID1414]. Appendices for Submission Addendum: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*: Janssen, 2022 [accessed 15.2.22]. 45p.

[3] Dold M, Bartova L, Kasper S. Treatment response of add-on esketamine nasal spray in resistant major depression in relation to add-on second-generation antipsychotic treatment. *Int J Neuropsychopharmacol* 2020;23(7):440-445.

[4] Byford S, Barrett B, Despiegel N, Wade A. Impact of treatment success on health service use and cost in depression: longitudinal database analysis. *Pharmacoeconomics* 2011;29(2):157-70.

[5] National Institute for Health Research (NIHR), Maudsley Biomedical Research Centre (BRC). *CRIS Natural Language Processing (NLP) Service: Library of production-ready applications (v1.6)[Internet]*. London: South London and Maudsley NHS Trust, 2021 [accessed 9.3.22] Available from: https://www.maudsleybrc.nihr.ac.uk/media/393677/applications-library-v16.pdf

[6] Wu B, Cai Q, Sheehan JJ, Benson C, Connolly N, Alphs L. An episode level evaluation of the treatment journey of patients with major depressive disorder and treatment-resistant depression. *PLoS One* 2019;14(8):e0220763.

[7] Talbot J, MacIntyre DJ, Rathod S, Morrens J, Young AH. Treatment resistant depression in the UK: sub-analysis of a European real-world evidence study. *BJPsych Open* 2021;7(Suppl 1):S55.

[8] Wolff R, Armstrong N, Ryder S, Fayter D, Worthy G, Noake C, et al. *Esketamine for treatment-resistant depression: ERG response to ACD*. York: Kleijnen Systematic Reviews Ltd, 2020 23p.

Please find my additional comments below on the assumed capacity of the mental health infrastructure to accommodate esketamine - "if" NICE gives approval for this treatment.

I hope this helps with any decisions about adjusting the 3-month funding mandate from NICE, but happy to discuss further if required.

As mentioned earlier - at this stage I simply don't have the capacity to undertake a detailed analysis of the readiness of all 54 MH trusts in England to adopt this technology, therefore I am taking a pragmatic "average case" scenario to estimate capacity.

1) Janssen pharmaceuticals advise 15,000 eligible patients per year – 54 MH trusts = 277 patients per trust (average) = 5 "NEW" patients per week per trust.
2) In line with SPC, treatment should be continued for 6 months after symptoms have improved
3) Treatment is twice a week for weeks 1-4, weekly for weeks 5-8, from week 9 onwards weekly or every 2 weeks,
4) Assume weekly treatment for the purpose of calculating capacity
5) Week 1 =5 "new" patients

On the basis of these assumptions - By Week 10 we will see each MH trust needing the infrastructure to manage 50 patients/week (5 new Patients and 45 follow up patients)

Assume capacity of a trust ECT facility allows 5 patients AM & 5 patients PM 5 days/week

Therefore to enable treatment for the target number of eligible patients additional infrastructure capacity is likely to be required beyond week 10 in order to continue to offer treatment "if" nice gives a positive opinion

There is uncertainty about number of patients dropping out of treatment and the proportion who will require weekly vs every 2 week visits. There is also uncertainty about the proportion of people who will require treatment to be continued beyond 6 month (after improvement) . The diversity of MH Trust size will mean that there will be variability in the actual time it takes for trusts to reach capacity to use existing infrastructure.

On the basis of the above I feel that most MH trusts will need a longer period of implementation than the usual 3 month mandate in order to put in place the required infrastructure to support the safe use of esketamine for the "average" number of 277 patients per trust per year.

By taking a gradual/iterative approach to treatment a 6 month implementation period should allow sufficient time for trusts to asses and then implement the infrastructure capacity required to support the supply/transport/storage/administration/post dose monitoring and waste disposal of esketamine nasal spray for the "real world" number of new and follow up patients from within their catchment area.

Whilst using the facilities & infrastructure within ECT facilities may offer the opportunity for most trusts to offer this treatment for a number of patients, relatively quickly – the downside for patients of using an ECT suite would be the need for additional transport/travel arrangements to be put in place. As you know the MH Long term plan highlights the importance of delivering care close to people's homes and within community settings.

Although there will be some loss of "economies of scale" by using/developing local community infrastructure as opposed to a centralised treatment "base" – the community option should be the desired long term goal. As previously advised MH trusts may need up

to 12months to plan – and implement the required community based infrastructure to support the supply/transport/storage/administration/post dose monitoring and waste disposal of esketamine within local community settings.

I hope this helps

With best wishes
Peter

Peter Pratt
GPhC 2023122
NHSE/NHSI National Speciality Advisor for Mental Health Pharmacy

Esketamine for Treatment-Resistant Depression

NHSE Comment

1. NHSE recognise there is a significant unmet need for treatment for people suffering from Major Depressive Disorder (MDD), including Treatment-Resistant Depression (TRD). Neither psychotherapy or current medications are effective for all people with TRD, residual symptoms persist, and treatment continuation does not always prevent relapse or recurrence of symptoms of depression.

2. Esketamine, the s-enantiomer of racemic Ketamine, is a non-selective and non-competitive N-methyl-D-aspartate (NMDA) receptor (NMDAR) antagonist that enhances glutamate release in the cortex and limbic system. Esketamine has a similar mechanism of action as Ketamine and target population with a higher affinity for the NMDA receptor compared with Ketamine. Esketamine is administered intranasally twice weekly for 4 weeks, once weekly for four weeks and beyond 8 weeks once weekly or once every 2-weeks, aiming for the least frequent dosing to maintain response.

3. NHSE note there is no consensus on what constitutes TRD. The company define TRD as people with MDD who fail to respond to two different oral antidepressants (OADs). The treatment pathway in NHS practice for TRD was recognised as being different to the populations included in the trials of Esketamine. Expert advice to committee indicated that in the NHS Esketamine would be used to treat TRD after failure of 3-4 OADs and after a trial of 1-2 augmentation therapies with an antipsychotic drug or lithium combination. (See later comments).

4. NHSE note Esketamine has a higher treatment burden than oral therapies, given it is a schedule 2 drug with a requirement for supervised administration and monitoring every week or two weeks for a 2-hour time period. Further, Esketamine pharmacology indicates its psychoactive effects play a role in both its therapeutic effect and its abuse potential. Adverse events that may be sought for abuse purpose (eg; dissociation, sedation, euphoric mood, hallucination, derealisation), clinical signs related to substance abuse disorder (eg; tolerance, withdrawal syndrome, drug dependence) and misuse (off-label use) have been identified in pharmacovigilance studies (Baudot et al).

   NHSE note post-marketing safety concerns have been reported following analysis of the FDA's Adverse Events Reporting System (Gastaldon et al). NHSE acknowledge the data only allow examination of what people on Esketamine are reporting (2,300 adverse events described by 960 people) and cannot address long-term issues of withdrawal and dependence. The authors controlled for confounding by indication by comparing the rate of a particular adverse event to the rate being reported for the oral antidepressant venlafaxine. The authors report adverse events not identified in the clinical trials, an increased risk of mania relative to Venlafaxine (a drug thought more likely to induce mania than other oral antidepressants) and a markedly increased risk of suicidal ideation. NHSE

note the safety profile of Esketamine in a real world population may be different from that described in the clinical trials.

NHSE accepts that physical effects for Esketamine prescribed for 8-weeks likely do not exist but concern remains about the safety and tolerability of Esketamine for the treatment of mood disorders; given the effects of chronic treatment and retreatment for variable periods are not known. It is currently unclear for how long and how often clinicians should prescribe the drug. It is unknown if problems will occur if dosing frequency is increased with loss of response or with discontinuation of Esketamine after long-term administration.

5. NHSE note the key clinical effectiveness evidence came from the TRANSFORM-2 and SUSTAIN-1 studies with supporting evidence for TRANSFORM-1 and TRANSFORM-3 and a long-term safety study SUSTAIN-2. The Montgomery-Asberg Depression Rating Scale (MADRS) measures severity of depression. Primary outcomes of response and remission in TRANSFORM-2 and relapse rates in SUSTAIN-1 were measured using MADRS. NHSE note the committee has discussed the strengths and weaknesses of the evidence base in detail previously.

   NHSE believes efficacy must be clearly established for a new and expensive therapy especially when there is uncertainty about the appropriate positioning of Esketamine in treatment algorithms, comparative effectiveness, potential for abuse or misuse, and appropriate setting, infrastructure and personnel required for competent and safe administration.

   There is new evidence relating to magnitude of treatment effect of Esketamine:

   In a meta-analysis of 5 trials (Papakostas et al) Esketamine + OAD decreased MADRS scores more than placebo + OAD (n=774; SMD -0.36; 95% CI -0.49 to -0.26). More recently, a Cochrane review (Dean et al) meta-analysis of 6 trials found Esketamine + OAD reduced MADRS score more than placebo + OAD at 4 weeks (N=1182; SMD -0.27; 95% CI -0.39 to -0.16).

   SMD refers to Standard Mean Difference, which is a measure of effect size for a continuous variable adjusted for scale and precision (SD) ie the point estimate of treatment effect. In terms of magnitude: SMD 0.2 is a small effect; SMD 0.5 is a medium effect and SMD 0.8 is a large effect.

   NHSE note there is debate regarding the change in MADRS score that represents a minimal clinically important difference (MCID). Regardless, NHSE note the evidence indicates the magnitude of the treatment effect to be modest with an effect size at the lower range of what may be recognised clinically with confidence intervals that straddle even the most conservative values of MCID.

   NHSE also note a more recent publication of intranasal Esketamine + OAD Vs Placebo + OAD in a Japanese population with TRD. At day 28 no statistical or clinically significant difference in change from baseline in MADRS score was apparent for any dose of Esketamine plus OAD

versus placebo plus OAD. The outcomes from this trial have not been included in the cited meta-analyses.

6. NHSE appreciate the efforts made by the company to position Esketamine in the treatment pathway (after failure of 3+ OADs or 3+ OADs and 1 augmentation with an antipsychotic or lithium); considered "optimised" populations with a high unmet need that may be more implementable in the NHS. NHSE note new evidence submitted by the company to support the use of Esketamine in the 2 optimised populations that are considered most likely to benefit from the intervention. NHSE understands this statement refers to the improved relative benefit of Esketamine/OAD Vs placebo/OAD, primarily because of the lower efficacy found with placebo/OAD in the 3+ OAD failures subgroup in TRANSFORM 2 compared with that recorded in the TRD population after failure of 2 OADs.

   NHSE have the following comments:

   NHSE accepts the cohorts who fail multiple courses of OADs and OADs plus 1 augmentation therapy represent a population more treatment resistant than the overall population in the TRANSFORM trials. NHSE note evidence presented by the company from the Discover dataset showing resource use with increasing line of therapy and treatment resistance (Table 2 company submission). Compared with the TRD population who failed 2 OADs, NHSE note the data shows a trend for increased resource use with line of therapy and treatment resistance. This trend was not apparent for elective and non-elective admissions. As all patients in the optimised populations are treated in secondary care, access may be problematic given the large catchment areas covered by mental health facilities. It is unclear if the optimised populations will require more frequent and higher doses of Esketamine to achieve response/remission and how long response/remission would be maintained, compared with the TRD population after failure of 2 OADs.

   NHSE note (Table 3 company submission) the optimised populations have higher rates of suicidal ideation/intent and comorbidities compared with the population after failure of 2 OADs. NHSE note people with suicidal ideation/intent and alcohol and substance abuse were excluded from the TRANSFORM trials. NHSE note the SmPC states the effectiveness of Esketamine in preventing suicide or in reducing suicidal ideation has not been established. NHSE note the SmPC for Esketamine states individuals with a history of drug abuse or dependence may be at greater risk for abuse and misuse and patients with a history of suicide related events or those exhibiting a significant degree of suicidal ideation should receive careful monitoring during treatment. NHSE note higher rates of associated comorbidities will alter risk/benefit considerations in the optimised populations and it is possible physical problems may be more apparent with long-term treatment and with drug discontinuation. NHSE note there is no consideration for costs and resource use explored in the model for treatment of people who may abuse or become dependent on Esketamine.

   NHSE note the estimates of response/remission rates for the optimised populations presented in Table 4 of the company submission. NHSE note the response/remission rates for Esketamine/OAD in the 3+OAD failures and after augmentation were calculated by

applying the relative treatment effect between SUSTAIN 2 3+ prior OAD and 3+ prior OAD and augmentation subgroups to the TRANSFORM (TRANSFORM 2 and 3) trials. NHSE note SUSTAIN 2 is a non-comparative safety trial and the relative treatment efficacy for the 3+ OAD failures was originally presented to committee using data only from the TRANSFORM 2 study.

NHSE note the direct entry population (86.2% Vs 13.8% transfers from other studies) recruited to SUSTAIN 2 differs from those recruited to the TRANSFORM studies eg: SUSTAIN 2 required a MADRS score >22 for entry to the trial (Vs 28 in TRANSFORM studies); the mean baseline MADRS in SUSTAIN 2 was 31.4 (Vs 37 in TRANSFORM 2); and 26.9% of the trial participants had a history of suicidal ideation in the 6 months prior to trial entry.

NHSE note patients from TRANSFORM 3 are now included in the economic model. NHSE note this was a small study and Esketamine/OAD did not achieve statistical significance for the primary endpoint compared with placebo/OAD. The response and remission rates were much lower than recorded in the other TRANSFORM trials, and in contrast to the TRANSFORM 1 and 2 trials no rapid effect of Esketamine was observed, with the FDA noting curve separation from the OAD/placebo arm only after 22 days of therapy.

NHSE note the ERG comment that augmentation would be an appropriate comparator for the 3+ OAD failures. NHSE agrees and given antipsychotic medications and lithium have different modes of action, it is plausible that people who fail 3+ OADs and 1 augmentation may receive a further trial of augmentation with a different mode of action than initially employed.

NHSE note the review of SUSTAIN 2 and SUSTAIN 3 adverse events provided by the ERG. NHSE would add that the SUSTAIN 2 publication reported that 14.5% of people who did not exhibit suicidal ideation at baseline, reported new suicidal ideation during the study.

NHSE note the detailed critique by the ERG of the company evidence relating to resource utilisation and healthcare costs, and issues relating to the treatment cap and subsequent treatment efficacy for treating people with TRD. NHSE also note concerns expressed by the ERG relating to model structure, uncertainty relating to treatment efficacy and magnitude of health state costs, and have presented two additional scenarios for committee to consider.

Professor Gary McVeigh
Clinical advisor NHSE

ISBN: 978-1-4731-4893-2