Esketamine for the Treatment of Treatment-Resistant Depression: Effectiveness and Value

Draft Evidence Report

March 21, 2019

Prepared for

ICER Staff and Consultants	The University of Illinois at Chicago College of Pharmacy's Center for Pharmacoepidemiology and Pharmacoeconomic Research*
Steven J. Atlas, MD, MPH Associate Professor of Medicine Harvard Medical School, Boston Director, Practice Based Research & Quality Improvement Division of General Internal Medicine Massachusetts General Hospital, Boston **Foluso Agboola, MBBS, MPH** Director, Evidence Synthesis Institute for Clinical and Economic Review **Katherine Fazioli** Senior Research Assistant Institute for Clinical and Economic Review **Varun Kumar, MBBS, MPH, MSc** Associate Director of Health Economics Institute for Clinical and Economic Review **Ellie Adair, MPA** Program Manager Institute for Clinical and Economic Review **David Rind, MD** Chief Medical Officer Institute for Clinical and Economic Review **Steve Pearson, MD, MSc** President Institute for Clinical and Economic Review	**Daniel R. Touchette, PharmD, MA** Professor of Pharmacy Assistant Director, Center for Pharmacoepidemiology and Pharmacoeconomic Research University of Illinois at Chicago **Nicole Boyer, PhD** Research Fellow University of Chicago **Brian Talon, PharmD** PhD Student University of Illinois at Chicago **Bob G. Schultz, PharmD** PhD Student and Research Fellow University of Illinois at Chicago **The role of the University of Illinois at Chicago College of Pharmacy's Center for Pharmacoepidemiology and Pharmacoeconomic Research is limited to the development of the cost-effectiveness model, and the resulting ICER reports do not necessarily represent the views of the UIC.*

Date of Publication: March 21, 2019

Steve Atlas served as the lead author for the report. Foluso Agboola led the systematic review and authorship of the comparative clinical effectiveness section in collaboration with Katherine Fazioli and Noemi Fluetsch. Varun Kumar was responsible for oversight of the cost-effectiveness analyses and developed the budget impact model. Noemi Fluetsch and Madeline O'Grady coauthored the section on coverage policies. David Rind and Steve Pearson provided methodologic guidance on the clinical and economic evaluations. The role of the UIC modeling group is limited to the development of the cost-effectiveness model, and the resulting ICER reports do not necessarily represent the views of UIC. None of the authors above disclosed any conflicts of interest.

About ICER

The Institute for Clinical and Economic Review (ICER) is an independent non-profit research organization that evaluates medical evidence and convenes public deliberative bodies to help stakeholders interpret and apply evidence to improve patient outcomes and control costs. Through all its work, ICER seeks to help create a future in which collaborative efforts to move evidence into action provide the foundation for a more effective, efficient, and just health care system. More information about ICER is available at http://www.icer-review.org.

The funding for this report comes from government grants and non-profit foundations, with the largest single funder being the Laura and John Arnold Foundation. No funding for this work comes from health insurers, pharmacy benefit managers, or life science companies. ICER receives approximately 19% of its overall revenue from these health industry organizations to run a separate Policy Summit program, with funding approximately equally split between insurers/PBMs and life science companies. Janssen is the only life science company relevant to this review that participates in this program. For a complete list of funders and for more information on ICER's support, please visit http://www.icer-review.org/about/support/.

About Midwest CEPAC

The Midwest Comparative Effectiveness Public Advisory Council (Midwest CEPAC) – a core program of ICER – provides a public venue in which the evidence on the effectiveness and value of health care services can be discussed with the input of all stakeholders. Midwest CEPAC seeks to help patients, clinicians, insurers, and policymakers interpret and use evidence to improve the quality and value of health care.

The Midwest CEPAC is an independent committee of medical evidence experts from across the Midwest, with a mix of practicing clinicians, methodologists, and leaders in patient engagement and advocacy. All Council members meet strict conflict of interest guidelines and are convened to discuss the evidence summarized in ICER reports and vote on the comparative clinical effectiveness and value of medical interventions. More information about Midwest CEPAC is available at https://icer-review.org/programs/midwest-cepac/.

The findings contained within this report are current as of the date of publication. Readers should be aware that new evidence may emerge following the publication of this report that could potentially influence the results. ICER may revisit its analyses in a formal update to this report in the future.

In the development of this report, ICER's researchers consulted with several clinical experts, patients, manufacturers and other stakeholders. The following clinical experts provided input that helped guide the ICER team as we shaped our scope and report. None of these individuals is responsible for the final contents of this report or should be assumed to support any part of this report, which is solely the work of the ICER team and its affiliated researchers.

For a complete list of stakeholders from whom we requested input, please visit: https://icer-review.org/material/trd-stakeholder-list/

Expert Reviewers

Cristina Cusin, MD
Assistant Professor in Psychiatry
Massachusetts General Hospital

Dr. Cusin served as site PI for an esketamine trial sponsored by Janssen.

Phyllis Foxworth
Vice President of Advocacy
Depression and Bipolar Support Alliance

No relevant conflicts of interest to disclose, defined as more than $10,000 in health care company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers.

William S. Gilmer, MD
Clinical Professor of Psychiatry and Behavioral Sciences
Northwestern University Feinberg School of Medicine

Dr. Gilmer has received consulting and speaker fee honorarium from Sunovion and Otsuka and owns equity in Organovo, Jounce, and Gilead Sciences.

Table of Contents

List of Acronyms Used in This Report

AHRQ	Agency for Health care Research and Quality
TRD	Treatment-resistant depression
MADRS	Montgomery-Åsberg Depression Rating Scale
HAM-D	Hamilton Rating Scale for Depression
MDD	Major depressive disorder
ECT	Electroconvulsive therapy
FDA	Food and Drug Administration
rTMS	Repetitive Transcranial magnetic stimulation
NMDA	N-methyl-D-aspartate
SNRI	Serotonin and norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
RCT	Randomized controlled trial
AD	Antidepressant
USPSTF	United States Preventive Services Task Force
TEAE	Treatment-emergent adverse event
REMS	Risk evaluation and mitigation strategy

1. Introduction

1.1 Background

Major depressive disorder (MDD) is a common psychiatric condition, with an estimated 16 million adults or 7% of adults in the United States experiencing at least one major depressive episode each year.[1] Symptoms of depression can include persistent sadness, feelings of hopelessness, loss of interest in usual activities, decreased energy, difficulty concentrating or sleeping, change in appetite and thoughts of hurting oneself. Depression can increase the risk of suicide and result in long-term suffering. It impacts all aspects of life including social relationships and the ability to work, and is the second leading cause of disability in the United States.[2] Treatment, including medication and psychotherapy, leads to improvement in many individuals, but multiple iterations in the therapeutic regimen may be required to achieve an adequate outcome. Treatment-resistant depression (TRD) refers to a major depressive episode with an inadequate response to therapy of adequate dosing and duration.[3,4] The failure of at least two trials of antidepressant monotherapies in the current episode is considered to indicate TRD,[5] but the number of trials has not been standardized.[6] Overall, approximately one in three patients with depression are considered "treatment-resistant." Patients with TRD have higher costs of care, decreased work productivity and account for around $64 billion in total costs.[3,7]

A major depressive episode is diagnosed based upon patient-reported symptoms of at least two weeks duration; there is a lack of reliable signs or tests that confirm the diagnosis or predict response to a specific treatment.[8] A diagnosis is typically made and treatment is often initiated by primary care clinicians, and broadly includes a range of different medications and psychological therapies in addition to supportive care such as self-help, relaxation techniques, and exercise. Second generation antidepressants including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), and atypical antidepressants (such as bupropion) are commonly used for initial pharmacotherapy in patients with depression.[9,10] However, patients with depression vary in terms of the severity of symptoms, course (episodic or chronic), and associated conditions such as anxiety or substance use disorders. Initial treatment may not work or may cause unacceptable side effects and switching to a different therapy is common. Since a trial of a therapy may require dose adjustments and six to 12 weeks to assess response, patients may find it difficult to remain on therapy long enough for an adequate trial of the treatment, especially if there are side effects or symptoms that are incapacitating. For this reason, TRD can be difficult to define because it includes not only the number of unique treatments tried, but whether the trials were considered adequate.

In efficacy trials, response to therapy is traditionally defined as a 50% or greater decrease in score from baseline on a depression rating scale.[6] However, many responders may continue to have

symptoms and impaired function, and improvement in functional outcomes can lag behind and are only modestly correlated with improvement in symptoms.[11] Remission, which refers to symptoms below a minimal level, is associated with improved quality of life and lower likelihood of relapse.[12,13] Initial treatment does not result in response in about one in three patients and remission in about two in three.[12] Even after four successive treatments, remission may not occur in one in three highlighting the great need for new therapies focused on those individuals with resistant depression. Treatment options for individuals with TRD broadly include modifying antidepressant therapy or augmenting existing therapies with non-antidepressant medications (such as atypical antipsychotics).[3] Modification of antidepressant therapy can take several forms: attempting to optimize existing treatment by maximizing the dose used, switching to a new treatment, or adding on to an existing therapy. There is limited evidence comparing these different strategies.[14] Among those with TRD, there are patients with highly resistant depression with symptoms over long periods of time, with many sequential treatment regimens, and inadequate responses and/or multiple relapses. These patients face chronic disability and account for a disproportionate cost of care.[7]

For these most difficult to treat patients, referred to as having refractory depression, other strategies such as electroconvulsive therapy (ECT) and transcranial magnetic stimulation (TMS) may be tried.[15,16] ECT has been shown to be useful in those with highly resistant depression.[17] However, ECT requires anesthetic sedation and has side effects including memory loss and cognitive impairment as well as major logistical constraints and stigma based upon media portrayals. Though patients can relapse after ECT, it can be administered chronically to maintain remission in certain patients. TMS is another device-based treatment for refractory depression. Repetitive TMS has been shown to improve depressive symptoms but may be less effective than ECT and also has logistical constraints that make long-term therapy difficult.[16,18] If not already tried, depression-focused psychotherapy may be added to pharmacotherapy, but is generally not considered stand-alone therapy for refractory depression.[19]

Intervention: Esketamine

Despite available treatments, there are many individuals who do not respond to multiple therapies for whom new treatment options are needed. One potential new target for therapy is the N-methyl-D-aspartate (NMDA) receptor.[20] Interest in agents that target this receptor has been driven by the observation that ketamine, an anesthetic, can transiently improve symptoms of depression.[21] Short-term studies have shown benefit, but this drug is usually administered intravenously and has side effects as well as the potential for abuse or diversion. A new agent, esketamine (Spravato™, Janssen), was approved on March 5, 2019 by the FDA for patients with TRD. Ketamine is a racemic mixture of two stereoisomers. Esketamine is the S-enantiomer, which binds with greater affinity to the NMDA receptor. It is a non-selective, non-competitive antagonist of the NMDA receptor and is being studied as a nasal spray for the treatment of adults with TRD.

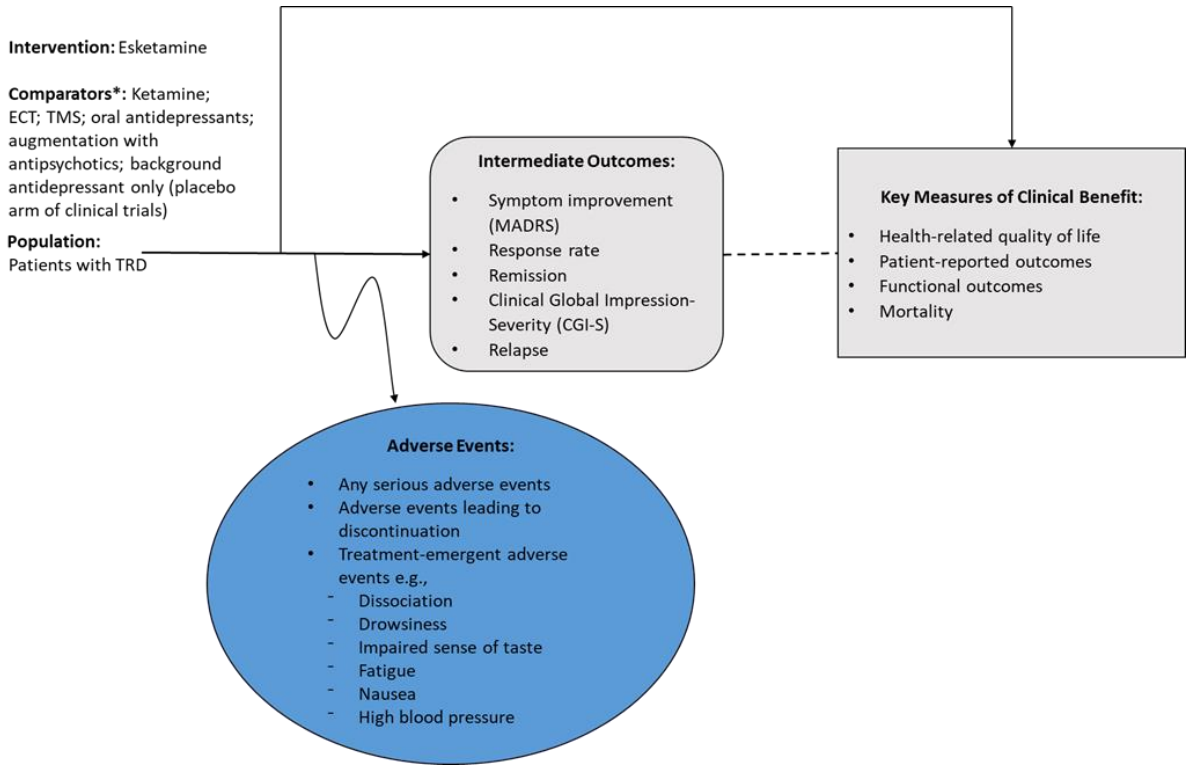
1.2 Scope of the Assessment

This review evaluated the comparative clinical effectiveness of esketamine for treatment-resistant major depressive disorder in adults. Evidence was collected from available randomized controlled trials, non-randomized clinical trials, comparative observational studies, as well as high-quality systematic reviews. We limited our review to those studies that captured the outcomes of interest. We included randomized controlled trials (RCTs) with at least 10 patients and sought evidence on esketamine and ketamine from non-randomized controlled trials and observational studies with at least 20 patients. We supplemented our review of published studies with data from conference proceedings, regulatory documents, information submitted by manufacturers, and other grey literature when the evidence met ICER standards (for more information, see https://icer-review.org/methodology/icers-methods/icer-value-assessment-framework/grey-literature-policy/). We sought head-to-head studies of esketamine and comparators to evaluate the feasibility of a network meta-analyses of selected outcomes.

Analytic Framework

The analytic framework for this assessment is depicted in Figure 1.

Figure 1.1 Analytic Framework



ECT: Electroconvulsive therapy, MADRS: Montgomery–Åsberg depression rating scales, TMS: Transcranial Magnetic Stimulation, TRD: Treatment-resistant depression
*Comparators may be used alone or in combination with background antidepressant.

The diagram begins with the population of interest on the left. Actions, such as treatment, are depicted with solid arrows which link the population to outcomes. For example, a treatment may be associated with specific clinical or health outcomes. Outcomes are listed in the shaded boxes: those within the rounded boxes are intermediate outcomes (e.g., change in blood pressure), and those within the squared-off boxes are key measures of clinical benefit (e.g., health-related quality of life). The key measures of clinical benefit are linked to intermediate outcomes via a dashed line, as the relationship between these two types of outcomes may not always be validated. Curved arrows lead to the adverse events of an action (typically treatment), which are listed within the blue ellipse.[22]

Populations

The population of focus for this review was adults ages 18 years and older with major depressive disorder, without psychotic features, and for whom two or more prior antidepressants prescribed at adequate dose and duration during the current episode have failed, termed TRD. We also sought evidence on key subgroups of patients suggested by patients and clinical experts. These included subgroups defined by:

- Age: Adults 18 – 64 years; Adults 65 years and older
- Number of prior treatment failures during the current episode (e.g., 2-3; 3-5; ≥5)

Interventions

The intervention of interest was esketamine nasal spray plus background antidepressants (continued or new administration). In addition, we sought clinical evidence on all forms of the product, including the intravenous form.

Comparators

Feedback from clinical experts suggested that esketamine will be used in patients for whom numerous antidepressants have failed. As such, our comparators for this review included treatments commonly used in this setting. These comparators may be used alone or in combination with background antidepressants (continued or new administration):

- Ketamine, an anesthetic agent used off-label for treatment-resistant depression
- ECT
- TMS

In addition, we sought evidence on the following comparators:

- Other oral antidepressants (plus background antidepressants)
- Augmentation with antipsychotics (plus background antidepressants)
- No additional therapy beyond background antidepressants (i.e., placebo arm of clinical trials)

Outcomes

We looked for evidence on the following outcomes of interest.

Efficacy Outcomes:

- Symptom improvement measured on Montgomery–Åsberg depression rating scales (MADRS) or other depression rating scale
- Rate of response
- Rate of remission
- Rate of relapse
- Symptom improvement as assessed by the clinician (Clinical Global Impression of Severity [CGI-S]) and patient (Patient Global Impression of Severity [PGI-S])
- Health-related quality of life assessed by EuroQol-5 Dimension-5 Level (EQ-5D-5L)

Safety Outcomes:

- Serious adverse events (including suicidality)
- Discontinuation due to adverse events
- Treatment-emergent adverse events (e.g.)
    - Dissociation
    - Dizziness
    - Headache
    - Fatigue
    - Somnolence
    - Nausea
    - Impaired sense of taste
    - High blood pressure
    - Metabolic changes
    - Substance use disorder
    - Memory loss

Timing

Evidence on intervention effectiveness and safety was derived from studies of at least fourteen days, as long as they met the study design criteria set forth above and measured an outcome of interest.

Settings

Evidence from all relevant settings was considered, including inpatient, outpatient/clinic, office, and home settings.

1.3 Definitions

Clinical Response, Remission and Relapse:

Outcomes of clinical trials of treatment of major depressive disorder commonly include response, remission and relapse. Clinical response is defined as at least a 50% reduction in the total score of an outcome measure. Clinical remission refers to a response that would be considered to result in symptoms that are absent or minimal. Remission will have a different cutoff depending on the measure, and there may be some differences for a given measure across different trials, as noted below. Finally, clinical relapse refers to recurrence of symptoms in one who has achieved a clinical response or remission. Patients achieving the definitions of response or remission and who remain in that state for a defined period of time are at risk for developing a relapse, or a new episode of MDD.

Hamilton Rating Scale for Depression (HRSD or HAM-D):

The HAM-D is the oldest and most widely used instrument to rate the severity of symptoms in depression. It was developed almost 60 years ago and was designed to assess the severity of depressive symptoms in hospitalized patients with melancholic type of depression.[23] It has been criticized in how it rates the various depressive symptoms, especially because it attributes higher weight to items of neurovegetative signs such as sleep and eating. The original 17-item questionnaire was later supplemented with 4 additional items that are generally not included in calculating a total score.[24] The first 17 items are typically included in a total score which ranges from 0 to 52 with 9 items rated in intensity or severity from 0 to 4 (0 = none/absent) and 8 symptom items rated from 0 to 2 (0 = none/absent).[25] Complete remission is generally considered to be a score of less than 7-10. It was designed to be administered by clinicians after a patient interview (either structured or unstructured). In addition, shorter and longer versions of the scale have been developed.

Montgomery-Åsberg Depression Rating Scale (MADRS):

The MADRS was developed to address some of the perceived short-comings of the HAM-D.[26] It provides a unidimensional assessment of the symptoms of depression with each symptom weighted similarly. It was derived from a 67-item scale,[27] and includes 10 items that showed response to treatment and correlated with the total score change. Individual items are rated in terms of severity from 0 to 6 (0 = no abnormality to 6 = severe), and complete remission is generally considered to be a score of less than 10-12. One study estimated the minimum clinically important difference (MCID) for MADRS to range from 1.6- to 1.9-point change from baseline.[28] Studies have also attempted to compare scores from the HAM-D with the MADRS.[29,30]

Nine-Item Patient Health Questionnaire (PHQ-9):

The PHQ-9 was originally designed to screen for depression in primary care and non-psychiatric settings and to track response to treatment.[31] The 9-item instrument is self-administered with the patient rating symptoms of depression in terms of severity from 0 to 3 (0 = not at all, 1 = several days, 2 = more than half the days, and 3 = nearly every day). The total score ranges from 0 to 27 with higher scores representing greater depressive symptoms.

Sheehan Disability Scale (SDS):

The SDS was developed in 1983 as a brief measure to assess functional impairment in three inter-related domains: work/school, social life/leisure activities, and family life/home responsibilities.[32] Three items assess how much symptoms have disrupted each of these domains on a 10-point visual analog scale (0 = not at all to 10 = extremely). These 3 items can be summed into a single dimensional measure of global functional impairment that ranges from 0 (unimpaired) to 30 (highly impaired).[33] Two additional questions ask about the number of days in the last week where symptoms led to lost or unproductive days at school or work.

1.4 Insights Gained from Discussions with Patients and Patient Groups

In developing and executing this report, we received valuable input from individual patients and patient advocacy groups throughout the scoping and evidence development process. We received public comments on our draft scoping document from three patient advocacy organizations: The National Alliance on Mental Illness, Mental Health America, and the Depression and Bipolar Support Alliance. We also conducted scoping calls with each of these organizations. Additionally, we conducted a focus group with three patients living with TRD. Below we summarize the key insights derived from this input.

Patients with treatment-resistant depression described different personal stories, but all had common themes that emphasized that MDD is a chronic disease that has profoundly affected all aspects of their lives and the lives of those close to them. Some reported excellent responses to prior therapies that subsequently waned over time, while others developed side effects that led them to have to stop therapy. In addition, some reported limited improvement with various therapies but never experienced full remission. The net result was that there was no single or combined therapy that offered them long-term control of their depressive symptoms.

Patients and patient advocacy groups highlighted the deficiencies with currently available treatments for depression. Despite a wide range of medications, both primary and adjunctive, used alone or in combination, many patients are unable to derive long-term benefit, either because they lose efficacy or develop intolerable side effects. Patients recognize that currently available therapies do not routinely provide long-term relief with minimal side effects. These side effects can

include metabolic changes resulting in weight gain, elevated blood sugar and cholesterol resulting in increased risk of diabetes, hypertension and vascular disease. They see this as reflecting insufficient knowledge about what causes depression to develop in the first place and then to persist over time. The focus on therapies that target a range of neurotransmitters is viewed as an advance, but not knowing which one to give for an individual patient and the recognition that some patients do not respond to therapies across available classes point to the need for increased support for basic research into the causes of MDD. Though depression-focused psychological therapies are commonly used, provide benefit and have fewer side effects than pharmacological therapies, they rarely are sufficient to control symptoms alone except in patients with milder forms of depression. Other non-medication therapies such as TMS and ECT have been shown to be effective, but also have high relapse rates, are time consuming and inconvenient, and especially for ECT may have cognitive side effects that make patients reluctant to consider treatment unless multiple other options have failed.

Patient advocacy organizations also raised systematic issues that they felt needed to be addressed. They highlighted that common outcome measures used in clinical literature may not adequately capture the impact of major depressive disorder on things that affect overall quality of life including relationships, work and family issues. They felt this to be particularly important for patients with treatment-resistant depression who were more likely to have severe symptoms over a long period of time and to have failed or not tolerated several prior therapies. Moreover, patients with MDD may have other psychiatric illnesses such as anxiety disorders that are impacted by depressive symptoms. Successful treatment of MDD may also help with these other conditions.

As a result, patients and patient advocacy groups suggest that symptoms of depression are more impactful on diminished quality of life than people realize. Stakeholders indicated that depression can be a serious and disabling condition that affects patients throughout their lives. When it occurs during formative educational years, it can prevent individuals from reaching their full academic potential, the result may be that measures of health-related quality of life used in economic analyses may not adequately reflect the true impact on those with treatment-resistant depression.

The toll of treatment-resistant depression also includes important economic costs. For some, the severity of symptoms and their duration prevent the ability to work at all. For others, the ability to work may be interrupted when symptoms flare or the nature of the treatment or its side effects may impact the ability to work. For example, some patients who derived benefit from IV ketamine reported they couldn't work full-time because of the time involved in going to an infusion clinic for therapy. Whether patients could not work at all, worked intermittently, part-time or were less productive at work because of symptoms of the depression or side effects of therapies, the net result was long-term under-employment with major socioeconomic impact.

Finally, some patients with treatment-resistant depression reported turning to off-label therapies through either their own investigation or at the suggestion of a clinician. We spoke with patients

who have tried ketamine, either IV or intranasally. For patients who reported benefit with ketamine, some expressed interest in the possibility of an FDA approved drug that is expected to work in a similar manner. Since ketamine is not covered by health insurers, patients commented on out of pocket costs that may be decreased if esketamine becomes covered by insurers. However, some worried that if esketamine was expensive, they still may have large out-of-pocket expenses through deductibles or non-coverage policies. In addition, they expressed concern about the time commitment to receive esketamine in a doctor's office. While it may be less than the time to receive IV ketamine, it would still require substantial time and inconvenience.

1.5 Potential Cost-Saving Measures in TRD

As described in its Final Value Assessment Framework for 2017-2019, ICER will now include in its reports information on wasteful or lower-value services in the same clinical area that could be reduced or eliminated to create headroom in health care budgets for higher-value innovative services (for more information, see https://icer-review.org/final-vaf-2017-2019/). These services are ones that would not be directly affected by esketamine (e.g., reduction in relapse), as these will be captured in the economic model. Rather, we are seeking services in current management of TRD beyond the potential offsets that arise from a new intervention. During stakeholder engagement and public comment periods, ICER encourages all stakeholders to suggest services (including treatments and mechanisms of care) currently used for patients with TRD that could be reduced, eliminated, or made more efficient.

2. Summary of Coverage Policies and Clinical Guidelines

2.1 Coverage Policies

To understand the insurance landscape for the treatment of TRD, we reviewed National and Local Coverage Determinations (NCDs and LCDs) from the Centers for Medicare and Medicaid Services (CMS), and publicly available coverage policies from representative public plans (Cigna HealthSpring, HealthNet, and WellCare) and national and regional private payers (Aetna, Anthem, and Blue Cross Blue Shield of Kansas City). We surveyed the coverage policies for esketamine, ketamine, electroconvulsive therapy (ECT), and repetitive transcranial magnetic stimulation (rTMS). No coverage policies for esketamine or ketamine were available at the time of this report: esketamine is awaiting FDA approval and ketamine does not have a label indication for treatment of depression.

We were unable to identify any NCDs or LCDs relating to the use of ECT for TRD. We only found one LCD, for the Wisconsin Physicians Service Insurance Corporation Medicaid plan, that provides coverage guidelines for the use of rTMS for treatment of severe MDD. While most public and private plans require failure of four antidepressants to necessitate treatment with rTMS, these guidelines state that rTMS may be indicated for patients who have failed at least one antidepressant in each of two separate classes during the current depressive episode.[34] Other local Medicaid plans surveyed, including Cigna HealthSpring, HealthNet, and WellCare, all outline medical policies for rTMS, but not ECT.[35-37] Cigna HealthSpring is the only policy that requires patients to have a documented diagnosis of TRD for treatments with rTMS; all other plans surveyed require a diagnosis of MDD.[35]

On the national level, both ECT and rTMS are covered as treatment options for TRD. The requirements for ECT treatment are very similar across national commercial plans: non-response to multiple pharmacotherapy trials of adequate dosage and duration, intolerance of effective medications due to side effects or medical counterindications, or a positive response to ECT treatment in previous depressive episodes is required by both Aetna and Anthem.[38,39]

Compared to ECT, prerequisites for rTMS treatment are more specific. Both Aetna and Anthem require the patient to have failed at least four antidepressants from at least two different classes in either the current or a previous episode.[40,41] In addition, Anthem requires the failure of two evidence-based augmentation therapies as well. Similarly, intolerance needs to be established by four trials of antidepressants with distinct side effects. Each treatment series with rTMS typically includes 36 sessions total (five days a week for six weeks and six tapering sessions over three weeks).[41] Blue Cross Blue Shield of Kansas City (BCBSKC) provides medical coverage guidelines for

rTMS but not ECT. As with the Aetna and Anthem plans, BCBSKC requires failure of at least four psychopharmacologic agents, as indicated by lack of significant improvement in depressive symptoms or inability to tolerate adverse events, for coverage of rTMS.[42]

2.2 Clinical Guidelines

Treatment recommendations have been developed by the American Psychiatric Association (APA), the Canadian Network for Mood and Anxiety Treatments (CANMAT), and the National Institute for Health and Care Excellence (NICE).[10,21,43,44] These guidelines cover a broad range of topics related to major depressive disorder and we summarize relevant issues pertaining to those with treatment-resistant depression.

American Psychiatric Association (APA)[21,43]

APA clinical practice guidelines for the treatment of major depressive disorder (MDD) do not specifically discuss "treatment-resistant depression" (TRD), but they describe strategies to address incomplete or nonresponse to treatment. The APA released the most updated guidelines for the treatment of patients with MDD in 2010.

Treatment for a major depressive episode may consist of pharmacotherapy, psychotherapy, somatic therapy (e.g., ECT or TMS), or the combination of two or more therapies. For patients whose response to pharmacological treatment of optimal dose and duration (typically four to six weeks) is incomplete, a change in treatment should be considered by the treating clinician. Several therapeutic options are available, such as switching to an antidepressant from the same pharmacological class (e.g., from one SSRI to another) or switching to a different pharmacological class of antidepressants (e.g., from an SSRI to a SNRI or a tricyclic antidepressant). Combination therapy with an antidepressant from another pharmacological class, or augmentation with a non-antidepressant medication (e.g., an antipsychotic or lithium) may also be considered as a next step. The guidelines also acknowledge that some patients might require doses that exceed than what is approved by the FDA to achieve therapeutic benefits.

ECT is recommended for patients who have either not responded to pharmacological or psychotherapeutic interventions or suffer from significant functional impairment. Treating clinicians are advised to consider ECT as a potential first-line treatment option for patients who have an urgent need for response (e.g., patients with severe MDD and at imminent risk of suicide) or in other instances where rapid antidepressant response is required. Furthermore, for patients who have a comorbid medical condition that would prevent the use of pharmacological therapies or have responded well to ECT treatment in the past, ECT should be considered as a treatment approach. Patient preference may also be factored in when considering ECT. Although TMS was approved by the FDA in 2008 for the treatment of MDD who had an insufficient response to at least one antidepressant trial in the current episode, clinical evidence for the use of TMS to treat MDD

was felt to be insufficient. For those who have responded to an acute course of ECT treatment, but not to pharmacotherapy, treatment with ECT may be continued during the continuation and maintenance phase of treatment.

In 2017, the APA released a consensus statement on the use of ketamine in the treatment of mood disorders.[21] Ketamine was noted to be beneficial for some patients, but they highlighted important limitations of the available evidence and potential risks. The statement emphasized the need for larger phase 3 trials with longer duration of treatment and follow-up but recognized that economic factors make it unlikely that such trials will be completed. Recommendations included establishing a registry of data from patients receiving ketamine in clinical practice. The World Health Organization also released a review of ketamine in 2015 that recognized its potential use as an antidepressant with a rapid onset of action.[45] The review noted ketamine's use in short-term trials and its potential for abuse. It assessed evidence for abuse world-wide and decided not to recommend bringing ketamine under international control as a drug of abuse.

Canadian Network for Mood and Anxiety Treatments (CANMAT)[10]

The CANMAT guidelines provide an algorithm to guide those with an inadequate response to an initial antidepressant, but these recommendations are also intended for those with treatment-resistant depression. The CANMAT guidelines highlight that consensus is lacking regarding the concept and definition of TRD. Even the common definition of inadequate response to 2 or more antidepressants does not take into account adjunctive strategies and those with varying levels of response. Effort should be made to ensure current treatment is optimized because of evidence showing that many patients receive subtherapeutic doses and/or inadequate duration of treatment. Options recommended broadly include switching to a second or third-line antidepressant versus adding an adjunctive agent. Because of limited evidence, the CANMAT guidelines emphasize an individualized approach based upon diagnostic reevaluation, consideration of previous medication trials, rational use of adjunctive medications, discontinuation of medications that have not been beneficial and careful monitoring. Ketamine was considered to be an experimental treatment and recommended use be limited to academic depression treatment centers. Finally, the CANMAT guidelines recommend that patients maintain treatment with antidepressants after achieving symptomatic remission for a variable time period based upon their risk for recurrence. For those with risk factors for recurrence, such as those with treatment-resistant depression, extending antidepressant treatment to 2 years or more is recommended.

Neurostimulation treatments were also considered in the CANMAT guidelines. We focus upon TMS and ECT here, though the guideline also considered vagus nerve stimulation (considered a third-line therapy), magnetic seizure therapy and deep brain stimulation (considered investigational only). TMS was considered to be a first line recommendation for patients with MDD who have failed at least 1 antidepressant. ECT was felt to remain a second-line treatment for patients with TRD, although it was considered first line in certain situations. Both TMS and ECT are often used as an

add-on to existing antidepressant regimens. Some evidence suggests starting TMS along with a new antidepressant is more effective than TMS alone. Despite limited evidence, TMS is thought to be less effective than ECT, particularly in patients who also have psychosis. However, fewer side effects are associated with TMS than with ECT. With both TMS and ECT, relapse is common without maintenance therapy.

National Institute for Health and Care Excellence (NICE)[44]

The clinical guidelines set forth by NICE were first published in 2009 and most recently updated in 2018. In its guidelines, NICE recommends that for patients who have not achieved satisfactory response to their initial antidepressant treatment of adequate dosage and duration, a change in treatment should be considered. Such options include adding psychotherapy to pharmacological treatments or switching antidepressants, either within the same pharmacological class or to a different class of antidepressants (e.g., from one SSRI to another or from an SSRI to a SNRI). Combination therapy with an antidepressant from another pharmacological class, or augmentation with a non-antidepressant medication (e.g., an antipsychotic or lithium) may also be considered as a next step.

ECT treatment may be suitable for the short-term treatment of individuals with severe MDD and at imminent risk of suicide, and when a rapid response is required, or when the patient has failed other treatments. NICE recommends against the use of ECT for the treatment of moderate depression unless the patient has not responded to multiple pharmacological and psychological treatments. In addition, the decision to use ECT should be made jointly with the patient as there are risks associated with this treatment modality. Continuous ECT treatment is only recommended for individuals who have previously responded well to ECT treatment and have failed other treatment options.

NICE has published interventional procedure guidelines for the use of repetitive transcranial magnetic stimulation (rTMS) for the treatment of MDD which were published in 2015. While the use of rTMS for the treatment of depression shows no major safety concern, clinical response may vary among patients. Nonetheless, NICE encourages clinicians to inform their patients about all possible treatment options, including rTMS, but also to reiterate that rTMS may not improve their depressive symptoms.

NICE is currently in the process of drafting guidelines for the use of esketamine for treatment-resistant depression.

3. Comparative Clinical Effectiveness

3.1 Overview

To inform our review of the comparative clinical effectiveness of esketamine for the treatment of patients with treatment-resistant depression (TRD), we abstracted evidence from available clinical studies of this agent, whether in published or unpublished form (e.g., conference abstracts or presentations, FDA review documents). As stated in the Background section, the comparators of interest were ketamine, ECT, TMS, oral antidepressants, augmentation with antipsychotics (e.g., olanzapine, aripiprazole, brexpiprazole, quetiapine) and no treatment beyond background antidepressants (i.e., placebo arms of clinical trials). Our review focused on clinical benefits (i.e., symptom improvement measured on MADRS or other depression rating scale; clinical response; remission; relapse; and health-related quality of life), as well as potential harms (drug-related adverse events).

3.2 Methods

Data Sources and Searches

Procedures for the systematic literature review assessing the evidence on esketamine for TRD followed established research methods.[46,47] We conducted the review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[48] The PRISMA guidelines include a checklist of 27 items, which are described further in Appendix Table A1.

We identified a previous systematic review of randomized control trials (RCTs) of ketamine, ECT, TMS, oral antidepressants, and augmentation for TRD which followed a similar scope to our review, with literature search end date of September 2014.[15] RCTs of ECT, TMS, oral antidepressants, and augmentation with antipsychotics that met our criteria from the systematic review were identified. In addition, we searched for new evidence that has emerged since 2014 by conducting an updated systematic literature search.

We searched MEDLINE, PsychINFO and EMBASE for relevant studies. The most recent search was conducted on December 5, 2018. In order to account for delays in indexing, the timeframe of our search for ECT, TMS, oral antidepressants, and augmentation with antipsychotics was overlapped with that of the previous systematic review, starting from January 2013 till December 2018. However, we conducted a de novo search for ketamine and esketamine till December 2018. Each search was limited to English-language studies of human subjects and excluded articles indexed as guidelines, letters, editorials, narrative reviews, case reports, or news items. We included abstracts from conference proceedings identified from the systematic literature search. All search strategies

were generated utilizing the Population, Intervention, Comparator, and Study Design elements described above. The proposed search strategies included a combination of indexing terms (MeSH terms in MEDLINE and EMTREE terms in EMBASE), as well as free-text terms.

To supplement the database searches, we performed manual checks of the reference lists of included trials and systematic reviews and invited key stakeholders to share references germane to the scope of this project. We also supplemented our review of published studies with data from conference proceedings, regulatory documents, information submitted by manufacturers, and other grey literature when the evidence met ICER standards (for more information, see http://icer-review.org/methodology/icers-methods/icer-value-assessment-framework/grey-literature-policy/).

Study Selection

After removal of duplicate citations, references went through two levels of screening at both the abstract and full-text levels. Three reviewers independently screened the titles and abstracts of all publications identified using DistillerSR (Evidence Partners, Ottawa, Canada) and disagreements were resolved through consensus.

Studies that did not meet the PICOTS criteria defined above, were excluded. Studies of oral antidepressants and augmentation with antipsychotics were only considered for inclusion if patients in the trial are also receiving background antidepressants. No study was excluded at abstract level screening due to insufficient information. Citations accepted during abstract-level screening were reviewed as full text. Reasons for exclusion were categorized according to the PICOTS elements.

Data Extraction and Quality Assessment

Two reviewers extracted data from the full set of included studies into an excel spreadsheet. Extracted data were independently verified by another researcher. Data elements included a description of patient populations, sample size, duration of follow-up, study design features (e.g., RCT or open-label), interventions (drug, dosage, frequency), outcome assessments (e.g., timing, definitions, and methods of assessment), results, and quality assessment for each study. We used criteria employed by the US Preventive Services Task Force (USPSTF) that included presence of comparable groups, non-differential loss to follow-up, use of blinding, clear definition of interventions and outcomes, and appropriate handling of missing data to assess the quality of clinical trials. For more information on data extraction and quality assessment, refer to Appendix D.

Assessment of Level of Certainty in Evidence

We used the ICER Evidence Rating Matrix to evaluate the level of certainty in the available evidence of a net health benefit among each of the interventions of focus (see Appendix D).[49]

Assessment of Bias

As part of our quality assessment, we evaluated the evidence base for the presence of potential publication bias. Given the emerging nature of the evidence base for newer treatments, we performed an assessment of publication bias for "esketamine" using the ClinicalTrials.gov database of trials. We scanned the site to identify studies completed more than two years ago that would have met our inclusion criteria and for which no findings have been published. Any such studies may indicate whether there is bias in the published literature. For this review, we did not find evidence of any study completed more than two years ago that has not subsequently been published.

Data Synthesis and Statistical Analyses

There were major differences in entry criteria, study populations, study design and outcome measures for clinical trials of esketamine versus other active treatments, so NMAs were not performed. Instead, we focused our attention on describing the comparisons made within the clinical trials of esketamine and its comparators, and where possible, we conducted random effect meta-analysis to combine data from multiple studies of esketamine.

3.3 Results

Study Selection

Our literature search identified a total of 2,342 potentially relevant references (see Appendix A Figure A1), of which five conference abstracts,[50-54] relating to five trials of esketamine (four Phase III RCTs and one open label trial) and two references,[55,56] relating to two trials of ketamine (One RCT & one single arm trial) met our inclusion criteria. We also considered evidence from 27 references relating to 13 RCTs of augmentation with antipsychotics (five RCTs of aripiprazole,[57-61] five RCTs of brexpiprazole,[62-66] one RCT of quetiapine,[67] and two RCTs of olanzapine[68,69]), 12 RCTs of TMS,[70-81] and one RCT of TMS & ECT[82] that met our inclusion criteria in order to assess the feasibility of NMA. Primary reasons for study exclusion during abstract and full text screening included use of interventions or comparators outside of our scope, wrong study population (e.g., MDD without TRD, active psychosis), small sample size (sample size < 10 for RCTs and < 20 for observational studies), minimum follow-up duration not met (at least 14 days), and conference abstracts with duplicate data as the full-text publications.

After further review of our included references, we noted that majority of the trials of augmentation with antipsychotics (five RCTs of aripiprazole[57-61], five RCTs of brexpiprazole[62-66], and the one RCT of quetiapine[67]) and one of the TMS trials[81] enrolled patients with less severe TRD compared to the esketamine trials. Specifically, these studies defined the cut-points for TRD differently, enrolled patients who had evidence of response to other antidepressants during pre-

randomization screening (such as between 25-50% on a depression rating scale) phase, while trials of esketamine used a screening criterion of less than 25% symptom reduction on MADRS scale. As such, the baseline depression severity in these trials differed significantly from the esketamine trials. Thus, we excluded these 13 trials from further consideration in our comparator evidence, and included and abstracted evidence from the remaining 14 trials (two trials of olanzapine[68,69], 11 trials of TMS[70-80], and one trial of TMS & ECT[82]) to further assess the feasibility of NMA.

Key Studies of Esketamine

Data to inform our assessment of esketamine were drawn from conference abstracts and supplemented by the FDA briefing document.

We identified four Phase III multicenter, RCTs of esketamine.[50-54] Three of them were similarly designed trials, two of which were conducted in patients 18 to 64 years of age (TRANSFORM-1 & -2),[50,51] while the third was conducted in patients aged 65years and older (TRANSFORM-3).[52] TRANSFORM-1 & -2 had similar inclusion criteria: patients with Diagnostic and Statistical Manual of Mental Disorders (DSM-5) diagnostic criteria for single-episode or recurrent MDD, without psychotic features, with a depression severity of 28 or more on MADRS scale, and non-response to one to five antidepressants in the current episode were eligible.[50,51] TRANSFORM-3 included patients with similar DSM-5 criteria, with a depression severity of 24 or more on MADRS scale, and non-response to one to eight antidepressants in the current episode.[52] Patients with psychotic symptoms or suicidal ideation with intent to act in the previous six months, or those that have had nonresponse to ECT or ketamine in the current episode were excluded from the trials.[50-52] Key trial characteristics is shown in Table 3.1.

All three trials included a four-week prospective screening and observational phase, in which patients continued the same oral antidepressants they were on in order to establish an additional failure, followed by a four-week randomized, placebo-controlled phase in which patients and investigators were blinded to treatment assignments (double-blind).[50-52] Patients who entered the double-blind phase must have had non-response (defined as ≤25% improvement) to at least two different antidepressant agents prescribed in adequate dosages for an adequate duration, with non-response to one antidepressant demonstrated in the prospective observational phase.[50-52] Patients were randomized to receive either esketamine or placebo nasal spray twice weekly, each combined with one of four choices of newly initiated open-label antidepressant (duloxetine, escitalopram, sertraline, or venlafaxine). Patients in the esketamine arm of TRANSFORM-1 received fixed doses of 56 mg or 84 mg,[50] while a flexible dosing schedule was used in TRANSFORM-2 & -3.[51,52] All three trials assessed the change from baseline in MADRS total score at week four as their primary efficacy outcome. Secondary outcomes included response rate (at least 50% improvement on MADRS score), clinical remission rate (MADRS≤12), early onset of sustained clinical response (≥50% reduction in MADRS on day 2 maintained through day 29), Clinical Global Impression-severity (CGI-S), and patient reported outcomes (Patient Health Questionnaire-9 [PHQ-9], Sheehan Disability

Scale [SDS]). The primary endpoint, assessment of MADRS, was conducted by remote, independent raters at 24 hours after the first dose, and weekly thereafter.

The fourth Phase III trial was a randomized withdrawal study that was designed to primarily assess relapse prevention (SUSTAIN-1).[53] SUSTAIN-1 enrolled patients either from TRANSFORM-1 or -2 or took direct entry patients who met the same inclusion and screening criteria as patients in TRANSFORM-1 & -2. Patients enrolled from TRANSFORM-1 & -2 must have completed the trial and demonstrated clinical response. The trial included a four-week induction period (for only direct enrolled patients), during which patients received twice weekly esketamine (56 mg or 84 mg) plus newly initiated oral antidepressant, followed by a 12-week optimization phase for responders, during which patients continued with the same dose of esketamine plus antidepressant at less frequent esketamine dosing (weekly for four weeks, then individualized to weekly or every other week based on symptoms), followed by a 48-week maintenance phase. In the maintenance phase, patients who were stable remitters (MADRS≤12 in at least three of four weekly assessment conducted in weeks 12-16) or stable responders (but were not in stable remission) were separately randomized to either continue with esketamine nasal spray plus oral antidepressant at current dose or switched to placebo plus oral antidepressant. The primary efficacy outcome was time to relapse in patients with stable remission. The key secondary outcome was the time to relapse in patients with stable response.

We also identified one open-label, long-term, multicenter, Phase III trial of esketamine (SUSTAIN-2) designed primarily to evaluate the long-term safety of esketamine.[54] SUSTAIN-2 enrolled patients from TRANSFORM-3 or took direct entry adult patients with single-episode or recurrent MDD, without psychotic features, with a depression severity of 22 or more on MADRS scale, and non-response to two or more antidepressants in the current episode. The trial consisted of a four-week screening phase (direct entry patient only), four-week induction phase (direct entry patient only and transferred nonresponders), 48-week maintenance phase (responders in induction phases only), and a four-week follow up phase.

Table 3.1. Phase III Randomized Trials of Esketamine

Phase III RCTs	Treatment Phases & Duration	Randomized Groups	Baseline Characteristics of Randomized Patients	Key Outcomes
TRANSFORM-1 Fixed Esketamine Dose **Adult 18-64 Years**	4-week prospective observation phase + 4-weeks RCT + 24-week follow-up	Esketamine 56 mg + AD Esketamine 86 mg + AD Placebo + AD	N=342 Mean age: 47 Current episode duration (yrs.): 3.9 MADRS mean: 37.5 Past failures of ≥ 3 ADs: 40%	MADRS change Clinical remission Clinical response
TRANSFORM-2 Flexible Esketamine Dose (56 mg or 84 mg) **Adult 18-64 Years**	4-week prospective observation phase + 4-weeks RCT + 24-week follow-up	Esketamine + AD Placebo + AD	N=223 Mean age: 46 Current episode duration (yrs.): 2.2 MADRS mean: 37 Past failures of ≥ 3 ADs: 36%	MADRS change Clinical remission Clinical response
TRANSFORM-3 Flexible Esketamine Dose (28 mg or 56 mg or 84 mg) **Adult ≥ 65 Years**	4-week prospective observation phase + 4-weeks RCT + 24-week follow-up	Esketamine + AD Placebo + AD	N=137 Mean age: 70 Current episode duration (yrs.): 4.1 MADRS mean: 35 Past failures of ≥3 ADs: 39%	MADRS change Clinical remission Clinical response
SUSTAIN-1 Flexible Dose (56 mg or 84 mg) **Adult 18-64 Years**	16-week open -label induction phase + 48-week (variable) randomized maintenance phase + 2-week follow-up	Esketamine + AD Placebo + AD	N=297 Mean age: 48 Current episode duration: NR Past AD failures: NR Stable remitters, MADRS mean: 37.5 Stable responders, MADRS mean: 39.5	Relapse

AD: antidepressant, MADRS: Montgomery-Åsberg Depression Rating Scale, N: number at randomization, NR: not reported, RCT: randomized controlled trial

Ketamine Studies

We identified one RCT of IV ketamine that met our inclusion criteria (Singh 2016).[55] Singh 2016 was a Phase II trial that enrolled adult patients with DSM-4 criteria for recurrent MDD, without psychotic features, who experienced an inadequate response to at least two antidepressants (at least one in the current episode). The trial consisted of a four-week double-blind, placebo-controlled phase followed by an optional two-week open-label treatment period. Sixty-eight eligible participants were equally randomized in the double-blind phase to one of four treatment arms: IV ketamine (0.5 mg/kg) twice- or thrice- weekly or IV placebo twice- or thrice-weekly. Participants in all arms continued their current antidepressant at the same stable doses for the duration of the study. At baseline, the mean age of patients enrolled was 44 years, the mean depression severity on MADRS scale was 35, and the majority of patients (85%) had failed only one or two antidepressants in the current episode. The mean duration of the current episode was not reported. The primary outcome was the change in MADRS from baseline to day 15 of the double-blind treatment period. Secondary endpoints included early onset of sustained clinical response (≥50% reduction in MADRS at week one maintained through day 15), clinical response rate (≥50% reduction in MADRS) at day 15, remission rate (MADRS≤10) at day 15, change in MADRS from baseline to day 29, and patient reported outcomes at day 29.

We also identified a single arm study of ketamine conducted in patients with chronic or recurrent MDD, who failed to respond to at least two FDA approved antidepressants in the current episode. [56] Murrough 2013 was designed to assess time to relapse among patients who achieved clinical response (≥ 50% improvement on MADRS scale) after a two-week course of IV ketamine. All patients were required to be off all antidepressants at the start of the trial (four weeks washout period for fluoxetine and two weeks for other antidepressants) and had to remain free of antidepressants throughout the treatment period. Following the two-week course of ketamine, ketamine was discontinued, and responders were followed twice weekly for 12 weeks or until relapse, which ever came sooner.

Other Comparator Studies

TMS & ECT

We reviewed 12 RCTs of rTMS that met our inclusion criteria, of which one was a head-to-head trial of rTMS versus ECT,[82] while the remaining 11 were sham-controlled trials.[70-80] In the head-to-head trial, 42 patients were randomized to either right unilateral ECT or unilateral rTMS. Of the 12 studies, eight were small, single-centered studies, conducted in different countries across the world.[73,74,76-80,82] The remaining four were larger, multicentered RCTs conducted in North America and Australia.[70-72,75] A majority of the studies enrolled patients with failure of two or more antidepressants in any episode (i.e., did not require failure in the current episode). At baseline, the average duration of current episode ranged from 0.8 (Pallanti 2010) to 3.6 years (Blumberger 2016);

and mean depression severity ranged from about 28 (Bakim 2012) to 38 (Rosa 2006) on the MADRS scale (for trials that assessed severity using only HAM-D, these scores were converted to MADRS score using the chart presented in Leucht 2018[30]). The mean number of past failures in the current episode was reported in only four of the studies, and it ranged from 1.3 to 3.5. Most of the studies assessed change in depressive symptoms, remission and response rates using a version of the Hamilton Rating Scale for Depression (HAM-D), while a few used the MADRS. Description of the study design, baseline characteristics of patients enrolled, and main efficacy outcomes observed in these trials are presented in Appendix Table D7.

Olanzapine

We identified two similarly designed studies of olanzapine conducted in adults with single episode or recurrent MDD, without psychotic features (Shelton 2005 & Corya 2006).[68,69] In Shelton 2005, patients were required to have a MADRS score of 20 or more at screening. Both trials required documented history of inadequate response to one SSRI plus an additional failure of an antidepressant during a seven-week prospective observational phase. In Shelton 2005, patients used nortriptyline during the prospective observational phase, while venlafaxine was used in Corya 2006. Following the prospective observational phase, patients who failed treatment (<30% improvement) were randomized to: continue the antidepressant taken during prospective phase plus placebo; olanzapine plus fluoxetine; olanzapine plus placebo; or fluoxetine plus placebo. Both trials had similar baseline characteristics (see Appendix Table D7). At baseline in Shelton 2005 and Corya 2006, the median duration of the current episode was approximately 12 months and 6 months, respectively, and the mean depression severity on the MADRS scale was 28 and 30, respectively. The mean number of past failures in the current episode was not reported in either trial. Both trials assessed changes in depressive symptoms, remission and response rate using the MADRS scale. Description of the study design, baseline characteristics of patients enrolled, and the main efficacy outcomes observed in both trials are presented in Appendix Table D7.

Comparability of Evidence Across Key Trials of Esketamine and Comparators

We considered conducting a network meta-analysis of two of the key clinical trials of esketamine that were homogenous in study populations, study design and outcome assessments (TRANSFORM-1 & -2),[50,51] the ketamine trial (Singh 2016),[55] the two trials of olanzapine (Shelton 2005 & Corya 2006), and the 12 trials of rTMS and ECT in order to quantitatively compare esketamine to the other interventions for TRD. However, key differences in entry criteria, study populations, study design and outcome measurements in these trials precluded these comparisons. The trials of esketamine, TRANSFORM-1 & -2, included patients with TRD, defined as patients with two or more failures of antidepressants in the current episode. However, as noted in the Background Section, the definition of TRD has not been standardized. As such, we found significant heterogeneity in how TRD was defined across trials, which was reflected in the differences in the inclusion criteria and the baseline characteristics of the patients in the trials. Many of the rTMS studies did not clarify

whether failures occurred in the "current" episode or during previous episode(s) (historical failure). The ketamine trial (Singh 2016) recruited patients with one or more failures in the current episode, while the olanzapine trials (Shelton 2005 & Corya 2006), although not explicitly stated, seemed to have included patients who prospectively failed only one antidepressant in the current episode. Another important difference noted was in the baseline MADRS severity. Trials of esketamine and ketamine seemed to have included patients with more severe depression (MADRS mean: 35 to 37) compared to some trials of olanzapine and rTMS. Finally, there were important differences in the design of the studies, such as the choice of using newly initiated concomitant antidepressant versus continuing a failed antidepressant; and in the definition of outcomes. These differences are summarized in Table 3.2.

Because of these differences, we did not think it was appropriate to perform a network meta-analysis across the trials. Instead, we focused on describing the comparisons made within the clinical trials of esketamine below and conducted a meta-analysis of two of the esketamine trials (TRANSFORM-1 & -2) that were homogenous in terms of inclusion and exclusion criteria, study design, and outcomes. Given that esketamine is the S-enantiomer of ketamine , we summarized the clinical benefit and harms in the trials of ketamine (see below). In addition, for context, we briefly summarized the clinical benefit and harms identified in the trials of olanzapine, TMS and ECT whose details are provided in Appendix Table D7.

Table 3.2. Comparability of Evidence Across Key Trials of Esketamine and Comparators

	Areas of Heterogeneity Among Clinical Trials	Esketamine	Ketamine	Olanzapine	TMS & ECT
Inclusion Criteria	Number of prior AD failures	≥2 in the current episode	At least 2 total failures, with ≥1 in the current episode	1 AD in the current episode & 1 historical failure	Majority of trials specified ≥2 historical and do not specify failure in the current episode
	Definition of failure	Non-response: ≤25% improvement in MADRS in a prospective phase	Inadequate response: definition not specified	Non-response: ≤30% improvement in MADRS in a prospective phase	Historical non-response
	MADRS severity	MADRS≥28	Not specified	MADRS≥20	Variable: ranges from MADRS≥20 to MADRS≥28*
Baseline Characteristics	Duration of current episodes, years	2.2 - 3.9	NR	Median: 0.5 - 1.0	0.7 – 4.0
	MADRS severity	37	35	28 - 30	28 to 38*
	Past failures of AD in the current episode	60% failed 1 or 2 at baseline; About 40% failed greater than 3 at baseline plus an additional prospective failure	About 85% failed only 1 or 2. No prospective failure	Failures in current episode NR	Failures in current episode 1.3 to 3.5 in four studies; NR in eight studies
Study Design	Concomitant AD	Newly initiated AD	Continued AD	Newly initiated AD	33% continued AD; 25% were not on AD; 42% mix of on and off AD
	Definition of remission	MADRS≤12	MADRS≤10	MADRS≤8	MADRS≤10 or HAM-D-17≤7

*For trials that assessed severity using only HAM-D, these scores were converted to MADRS score using the chart presented in Leucht 2018[30]

AD: antidepressant, HAM-D-17: Hamilton Depression Rating Scale, 17-item, MADRS: Montgomery-Åsberg Depression Rating Scale

Quality of Individual Studies

We did not assign an overall quality rating to any of the esketamine trials because all the references were obtained from grey literature sources (e.g., conference proceedings, FDA briefing documents). However, we highlighted the information available on each trial regarding the comparability of groups, participant blinding, intervention definitions, outcome definition, outcome reporting, and intention to treat analysis in Appendix Table D1. We noted some differential loss to follow up in TRANSFORM-1, with loss of 16 patients in the group taking 86 mg esketamine, compared to four and five patients in the 56 mg esketamine and placebo groups, respectively.

Clinical Benefits of Esketamine

Symptom Improvement, Clinical Response and Remission

In two Phase III trials conducted in adults (ages 18 to 64 years), symptom improvement at four weeks on the MADRS scale was greater with esketamine than placebo (all patients also received a new background antidepressant). A greater proportion of patients also achieved clinical response but not clinical remission with esketamine at four weeks.

In one Phase III trial conducted in adults ages 65 and older, symptom improvement at four weeks was not significantly different between esketamine and placebo (all patients also received a new background antidepressant).

The primary outcome in the RCTs of esketamine was improvement in symptoms, based on change from baseline in MADRS score at week four.[50-52] Clinical response, defined as at least 50% improvement in MADRS scale at week four from baseline; and clinical remission rate, defined as reaching 12 or less on MADRS scale at week four were secondary outcomes reported in these trials.

In TRANSFORM-2, flexible dosed esketamine plus antidepressant resulted in greater improvement in MADRS score compared to placebo plus antidepressant at four weeks (mean change from baseline (CFB) -21.4 vs. -17.0; least square mean difference [LSMD] -4.0; 95% CI: –7.31, –0.64; P =0.020) (Table 3.3).[51] In TRANSFORM-1, both doses of esketamine (56 mg and 84 mg) showed a numerically greater improvement from baseline compared to placebo (mean CFB -19.0 & -18.8 vs. -14.8), however, statistical significance was not demonstrated with the 84 mg esketamine plus antidepressant versus placebo plus antidepressant.[50] Therefore, the 54 mg dose was not formally evaluated based on predefined testing sequence. However, exploratory analysis showed that patients in the 56 mg arm of esketamine experienced a greater improvement compared to the placebo arm (LSMD -4.1;95% CI: -7.67, -0.49; p=0.0114).[83] As noted above, we conducted random effect meta-analysis of TRANSFORM-1 & -2. We pooled the two esketamine doses in TRANSFORM-1 (56 mg and 84 mg) into one single esketamine arm for the meta-analysis. Results of the meta-analysis was in favor of esketamine, showing a greater improvement on MADRS score for

esketamine plus antidepressant compared to placebo plus antidepressant (Mean difference: -3.84; 95% CI: -6.29, -1.39)(Figure 3.1).

A greater proportion of patients achieved clinical response and remission at four weeks in the esketamine arms compared to placebo in TRANSFORM-1 & -2, although statistical significance was not reported (Table 3.4). Meta-analysis of the two trials showed that compared to placebo plus antidepressant, patients on esketamine plus antidepressant were more likely to achieve clinical response (Relative risk [RR] 1.30; 95% CI: 1.08, 1.56) (Figure 3.3); the relative likelihood of clinical remission was similar but was not statistically significant (RR 1.37; 95% CI: 0.99, 1.91) (Figure 3.2). An additional secondary outcome related to clinical response was the proportion of patients showing onset of clinical response by day 2 that was maintained through day 28. Numerically more patients on esketamine plus antidepressant achieved early onset of sustained clinical response by day 2 (7.9% vs. 4.6%) in TRANSFORM-2, although the difference was not statistically significant. This outcome was not formally evaluated in TRANSFORM-1, however post-hoc analysis favored patients receiving both doses of esketamine compared to placebo treated patients (56 mg: 10.4% vs 1.8% [odds ratio [OR]: 6.5; 95% CI: 1.4, 60.5]; 84 mg: 8.8% vs 1.8%, [OR: 5.3; 95% CI: 1.1, 50.9]).[84]

In the study conducted in adults 65 years and older that included a flexible dosing schedule (TRANSFORM-3), patients on esketamine plus antidepressant also experienced numerically greater improvement on the MADRS scale compared to those on placebo plus antidepressant at four weeks (mean CFB –10.0 vs –6.3), however this was not statistically significant (Table 3.3).[52] Similar to the adult population aged 18-64, a greater proportion of elderly patients in the esketamine arm of the TRANSFORM-3 trial also achieved clinical response (23.6% vs. 12.3%) and clinical remission (15.3% vs. 6.2%) (statistical significance not reported).

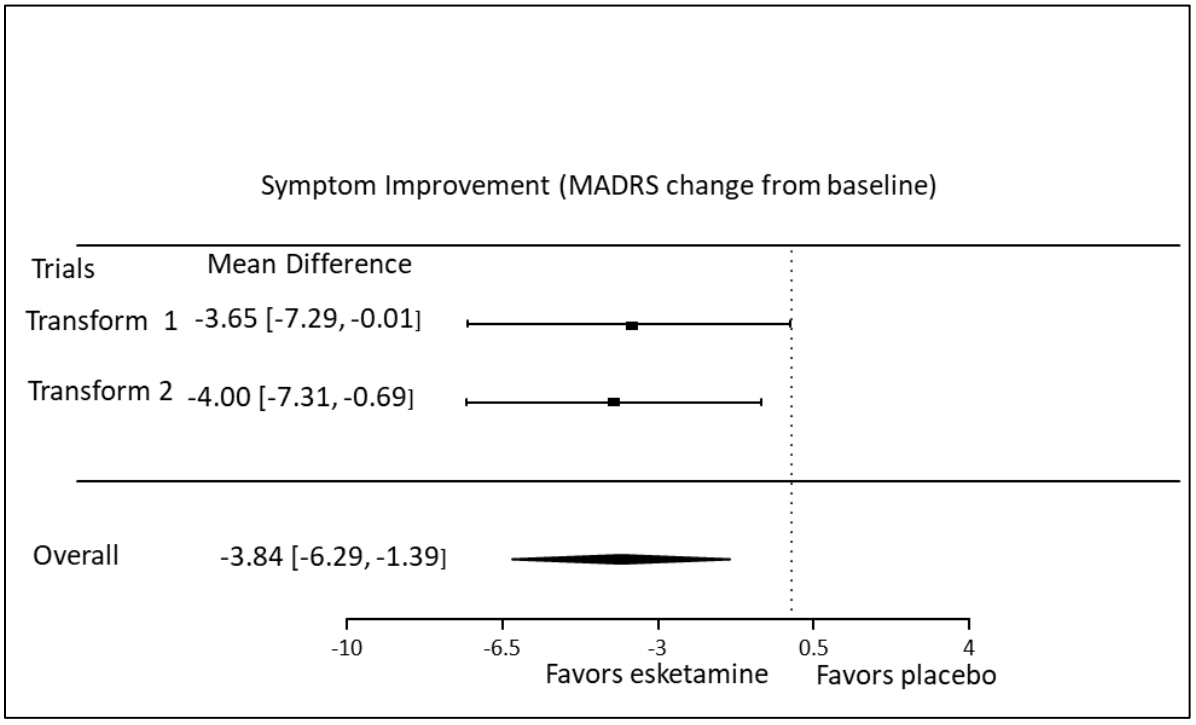
Table 3.3. Esketamine: Change on MADRS Scale Between Baseline and Four-Week Follow-Up

Trial	Intervention	N	Baseline (SD)	Δ (SD)	LS Mean Δ (95% CI)*	P-Value
Adult (18 to 64 years)						
TRANSFORM-1	Placebo + AD	113	37.5 (6.2)	-14.8 (15.1)	reference	---
	Esketamine 84 mg+ AD	114	37.8 (5.6)	-18.8 (14.1)	-3.2 (-6.88, 0.45)	0.088
	Esketamine 56 mg + AD	115	37.4 (4.8)	-19.0 (13.9)	-4.1 (-7.67, -0.49)	0.011
TRANSFORM-2	Placebo + AD	109	37.3 (5.7)	-17.0 (13.9)	reference	---
	Esketamine + AD	114	37.0 (5.7)	-21.4 (12.3)	- 4.0 (-7.31, -0.68)	0.020
Elderly (≥65 years)						
TRANSFORM-3	Placebo + AD	65	34.8 (6.4)	-6.3 (8.9)	reference	---
	Esketamine + AD	72	35.5 (5.9)	-10.0 (12.7)	-3.6 (-7.2, 0.07)	0.059

AD: antidepressant, CI: confidence interval, LS: least square, MADRS: Montgomery-Åsberg Depression Rating Scale, N: number analyzed, SD: standard deviation, Δ: change

*Least square mean difference estimated using mixed model for repeated measures (MMRM)

Figure 3.1. Esketamine Versus Placebo: Meta-Analysis of TRANSFORM-1 & -2: MADRS Change



CI: confidence interval, MADRS: Montgomery-Åsberg Depression Rating Scale

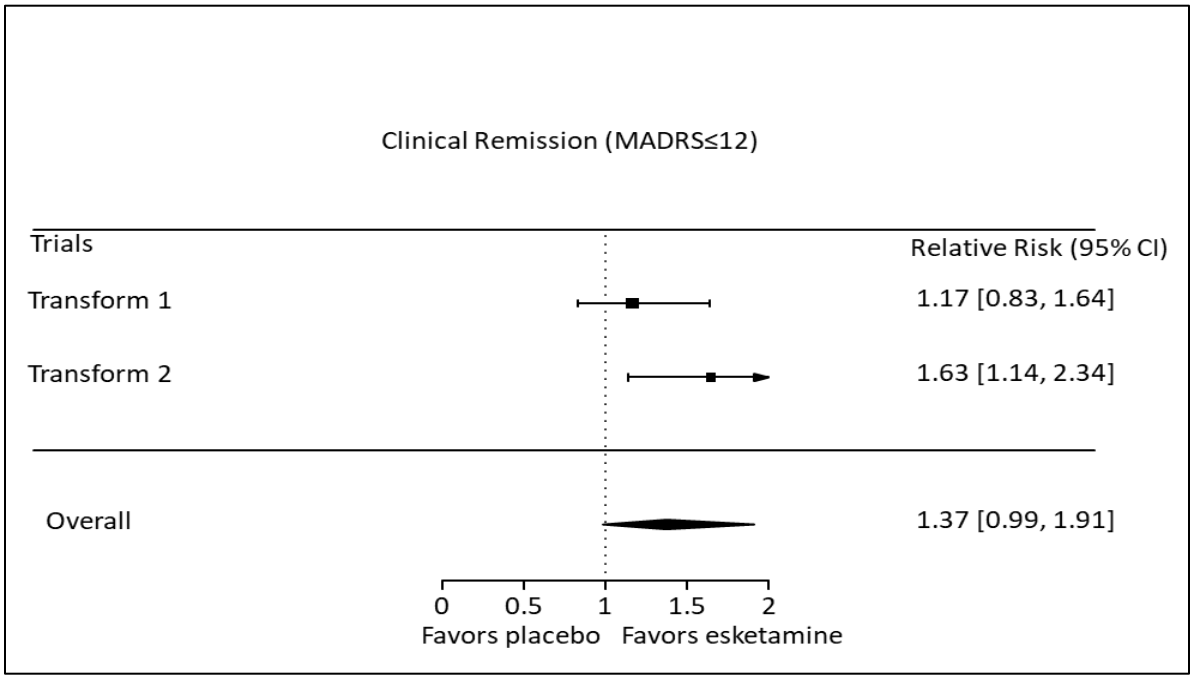
Random effects meta-analysis; I-squared: 0 %

Table 3.4. Esketamine: Clinical Response and Remission Based on MADRS Scale*

Trial	Intervention	N	Clinical Response (≥ 50% Improvement) Rate, %	Clinical Remission (MADRS ≤ 12) Rate, %
Adult (18 to 64 Years)				
TRANSFORM-1	Placebo + AD	113	37.2	29.3
	Esketamine 84 mg+ AD	114	45.2	33.3
	Esketamine 56 mg + AD	115	52.2	34.8
TRANSFORM-2	Placebo + AD	109	47.7	28.4
	Esketamine + AD	114	61.4	46.5
Elderly (≥ 65 Years)				
TRANSFORM-3	Placebo + AD	65	12.3	6.2
	Esketamine + AD	72	26.3	15.3

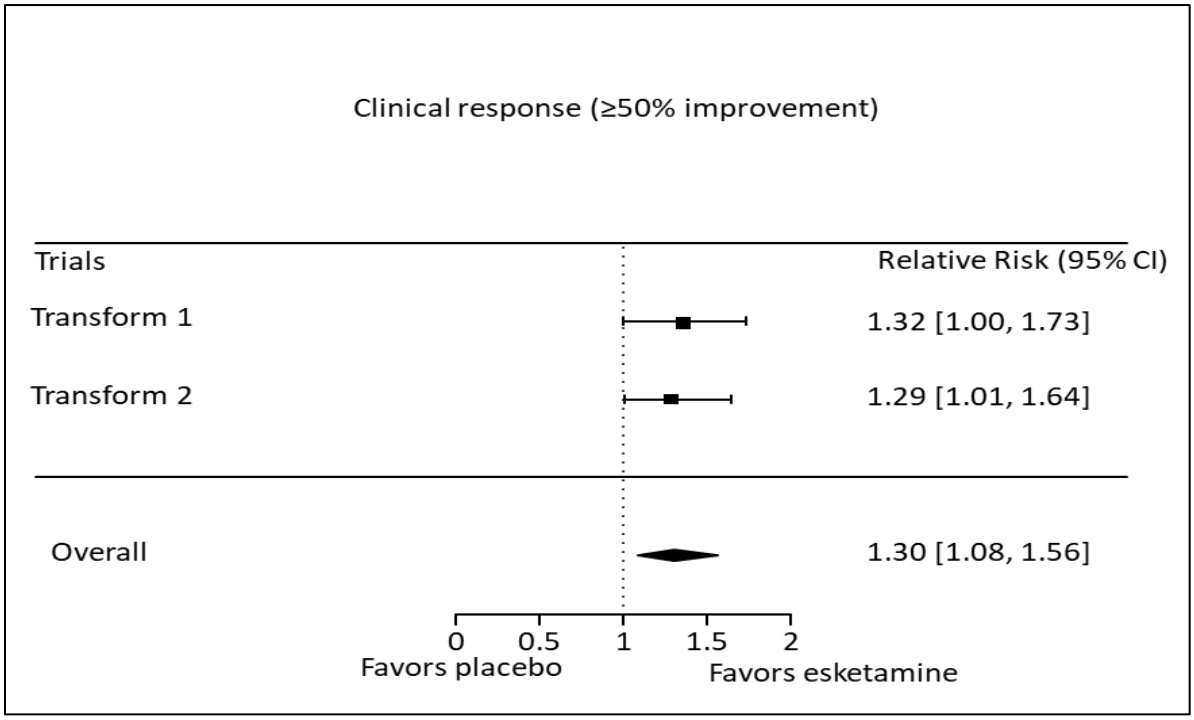
AD: antidepressant, MADRS: Montgomery-Åsberg Depression Rating Scale, N: number analyzed
*statistical significance not reported

Figure 3.2. Esketamine Versus Placebo: Meta-Analysis of TRANSFORM-1 & -2: Remission



CI: confidence interval, MADRS: Montgomery-Åsberg Depression Rating Scale
Random effects meta-analysis: I-squared 0 %

Figure 3.3. Esketamine Versus Placebo: Meta-Analysis of TRANSFORM-1 & -2: Response



CI: confidence interval, MADRS: Montgomery-Åsberg Depression Rating Scale
Random effects meta-analysis; I-squared 0 %

Relapse Prevention

In one randomized trial in adults who achieved stable clinical remission or stable clinical response, continued treatment with esketamine reduced the risk of relapse.

As described above, SUSTAIN-1 evaluated the time to relapse among stable remitters and stable responders who were randomized to either continue maintenance esketamine plus oral antidepressant or switch to placebo plus oral antidepressant for 48 weeks or until relapse, whichever came first. Stable remission was defined as achieving MADRS ≤ 12 for at least three out of the last four weeks of the 12 weeks optimization phase of receiving esketamine, while stable response was defined as achieving ≥ 50% reduction in MADRS total score from baseline in each of the last two weeks of the optimization phase, but without meeting criteria for stable remission. Relapse during the was defined as having a MADRS score of 22 or greater at two consecutive assessments and/or undergoing hospitalization for worsening depression, suicide attempt, suicide, or any other clinical event suggestive of relapse (as decided by investigators).[53]

Out of the 705 patients enrolled in SUSTAIN-1, 176 patients achieved stable remission, while an additional 121 patients only achieved stable response.[53] The median exposure to esketamine was 17.7 weeks versus 10.2 weeks for placebo among the stable remitters, while it was 19.4 weeks for esketamine versus 10.1 weeks for placebo among stable responders. Among the stable remitters,

26.7% of patients on maintenance esketamine plus antidepressant experienced a relapse compared to 45.3% among patients switched to placebo plus antidepressant.[53] Among the stable responders, 25.8% of patients on esketamine plus antidepressant experienced a relapse compared to 57.6% among patients switched to placebo plus antidepressant.[53] Time to relapse was statistically significantly delayed for stable remitter patients on esketamine compared to patients on placebo (p=0.003, Table 3.5).[53] Similarly, among the stable responders, time to relapse was in favor of esketamine plus oral antidepressant (p<0.001, Table 3.5).[53]

Of note, the FDA review committee noted that there was a faster rate of relapse observed in SUSTAIN-1 compared to other maintenance of effect studies of MDD. This could reflect functional unblinding, with patients on placebo realizing that they are no longer on esketamine after switching, given the immediate side effects associated with esketamine use.[83] However, there is insufficient evidence to support or reject this possibility. Overall, continued treatment with esketamine plus oral antidepressant maintenance dose in this trial decreased the risk of relapse by 51% among stable remitters (hazard ratio [HR] 0.49; 95%CI: 0.26, 0.84) and by 70% among stable responders (HR 0.30; 95% CI: 0.16, 0.55).[53]

Table 3.5. Time to Relapse

Trial	Randomized Patients	Interventions	Median Days to Relapse (95% CI)	Hazard Ratio (95% CI)
SUSTAIN-1	Stable remitters (N=176)	Placebo	273 (97, NE)	reference
		Esketamine	NE	0.49 (0.26, 0.84)
	Stable Responders (N=121)	Placebo	88 (46, 196)	reference
		Esketamine	635 (264, 635)	0.30 (0.16, 0.55)

CI: confidence interval; N: number analyzed; NE: not estimable

Patient-Reported Outcomes

Esketamine improved depressive symptoms as measured by patient health questionnaire-9; and improved quality of life as measured by Sheehan disability scale

Change from baseline on the patient health questionnaire-9 (PHQ-9) and Sheehan disability scale (SDS) were measured as secondary outcomes in esketamine trials. PHQ-9 is a 9-item patient reported instrument used to assess depressive symptoms on a scale of 0 to 27, with higher scores representing greater depressive symptoms. A change of five points in the PHQ-9 has been previously defined as the minimum clinically important difference (MCID).[85] SDS is a 5-item patient reported instrument used to assess functional impairment in work/school, social life, and family life on a scale of 0 to 30, with higher score representing greater impairment. MCID for SDS has not been previously specified.

In TRANSFORM-2, clinically significant reduction was observed in the PHQ-9 score for both arms of the trial, however, esketamine plus antidepressant resulted in greater improvement from baseline on PHQ-9 compared to placebo plus antidepressant at four weeks (mean change from baseline: -13.0 vs. –10.2; LSMD -2.4; 95% CI: –4.18, –0.69; p<0.006).[51]. Similarly, changes on SDS score favored esketamine plus antidepressant compared to placebo plus antidepressant at four weeks (mean change from baseline -13.6 vs. –9.4; LSMD -4.0; 95% CI: –6.28, –1.64; p<0.001), however, clinical significance of this change is not known.[51] Similar trends of greater improvement on PHQ-9 and SDS in favor of esketamine were also observed in TRANSFORM-1 & -3 trials (Appendix Table D4).

Harms

Adverse events with esketamine were mostly mild to moderate and resolved on dosing days. The most common were nausea, dissociation, and dizziness. Patients receiving esketamine were more likely to experience sedation, have clinically important increases in systolic and diastolic blood pressure, and discontinue treatment.

Overall, there were no new safety concerns reported in patients treated with esketamine for up to one year, and no evidence of increased risk of abuse/misuse was reported.

Most treatment-emergent adverse events (TEAEs), defined as those first reported or worsening in severity after initiating study treatment, in the placebo-controlled trials of esketamine were of mild to moderate severity.[84] The most commonly reported TEAEs, with incidence ≥5% and greater occurrence in the esketamine arm included nausea/vomiting, dissociation, dizziness, headache, vertigo, dysgeusia (distortion of sense of taste), somnolence, sedation, insomnia, blurry vision, increased blood pressure, paresthesia, hypoesthesia (reduced sense of touch or sensation), and fatigue (Table 3.6 and Appendix Table D6).[84] Most TEAEs occurred at a higher incidence in patients aged 18-64 years (TRANSFORM-1 & -2) than in patients aged 65 years and older (TRANSFORM-3), with the exception of increased blood pressure and fatigue.[84] In the fixed-dose study (TRANSFORM-1), rates of TEAEs were generally similar for the 56 mg and 84 mg dose.[84]

Eighty-six percent of TEAEs in the Phase III RCTs occurred on the day of intranasal medication administration, and majority of these events resolved on the same day.[84] Primary safety concerns occurring on the same day in a considerable higher proportion of esketamine treated patients compared to the placebo treated patients included dissociation, sedation, and increased blood pressure. Due to the high relative incidence of dissociation and sedation associated with esketamine as evidenced in the placebo-controlled trials, the FDA label for esketamine includes a boxed warning for sedation and dissociation, and states that patients should be monitored for at least two hours after administration.[86] The FDA label also includes a warning for increased blood pressure and notes that patients' blood pressure should be monitored pre- and post-dose, and the benefit versus risk of esketamine should be considered in patients for whom an increase in blood

pressure poses considerable risk.[86] See further details on evaluation of dissociation, sedation and increased blood pressure below.

Overall, the incidence of serious adverse events (SAEs) in the short-term esketamine trials was low (<5%) in both the esketamine- and placebo-treated groups (Table 3.6).[84] There were no apparent differences in the rates of most SAEs between the esketamine and placebo groups, with the exception of SAEs of depression and suicidal ideation occurring at a higher rate in the esketamine arms in TRANSFORM-1.[83] Discontinuation due to AEs were higher among the esketamine-treated patients compared to the placebo-treated patients (Table 3.6).[84]

We identified one long-term, open-label study that evaluated the safety of esketamine dosed weekly or every other week for up to 48 weeks in 603 patients who responded to esketamine during a four-week induction phase (SUSTAIN-2). During the trial, 24% of patients received weekly dosing throughout, 38% changed from once weekly to every other week dosing, and 38% changed back and forth from weekly and every other week dosing.[54] The study was terminated when the predefined exposure criteria were met (at least 300 patients with six months exposure and at least 100 patients with 12 months exposure). Eighty-six percent of participants reported at least one TEAE during the 48-week maintenance phase, most of which occurred on dosing days and resolved on the same day.[54] The most common TEAEs were generally similar to those reported in the short-term esketamine trials (Appendix Table D6).[54] SAEs were reported in about 6% of patients, and the most commonly reported SAEs included depression, suicidal ideation, suicide attempt, and gastroenteritis[54] In all, about 10% of participants discontinued esketamine due to TEAEs, with more patients discontinuing treatment during the induction phase (6.8%) compared to the maintenance phase (3.8%) (Appendix Table D6).[54]

Data from the placebo controlled trials and the long-term, open-label study did not show an increased risk of interstitial cystitis, liver injury, or impaired cognitive function in esketamine-treated patients, all of which are commonly-reported complications associated with repeated use of ketamine.[83] Patients 65 and older did experience a slowing of reaction time during the long-term safety study (SUSTAIN-2), but there was insufficient data to support that the effect was due to esketamine.[84] In addition, there was no evidence of drug-seeking behavior or misuse or abuse of esketamine in any of the trials,[84] although the details of how this was assessed are not clear. However, the FDA label includes a boxed warning for abuse and misuse due to its similar pharmacological profile to ketamine, confirmed in a human abuse potential Phase I study.[86] Furthermore, a Risk Evaluation and Mitigation Strategy (REMS) has been put in place for the use of esketamine due to the concerns around dissociation, sedation, and misuse and abuse (ketamine is misused and abused for its dissociative and hallucinogenic effects).[86] REMS is a drug safety program that the FDA has the authority to require for medications with serious safety concerns to help ensure that the benefits of the medication outweigh its risks. [83]

A total of six deaths occurred during the esketamine development program (five during the Phase III trials, and one during the Phase II trial), all in esketamine-treated patients, although none was considered by the investigators to be esketamine-related.[83] Three deaths were by suicide, occurring after the patient's last dose of esketamine; one death was from a motorcycle accident (occurred 26 hours after esketamine use, therefore unlikely to be sedation related); one was a sudden death in a 60-year old patient with hypertension and obesity (all vitals were normal during patient's visit 5 days prior to death); and one was myocardial infarction in a 74-year old patient with history of hypertension and hyperlipidemia (occurred 6 days after last dose of esketamine).

Dissociative Symptoms

Dissociation was the one of the most commonly reported treatment emergent adverse event associated with the use of esketamine and was generally reported using different terms such as 'spacey', 'sense of floating', and 'feeling of faintness'. The Clinician-Administered Dissociative States Scale (CADSS) was used to objectively assess present state dissociative symptoms and transient perpetual effect pre-dose, and 40- and 90-minutes post-dose during the clinical trials of esketamine. CADSS scores range between zero and 92, and scores between zero and four are considered normal.[84] Across the three short-term trials, 60% to 79% of patients receiving esketamine experienced more than a four point increase in CADSS scores following dose administration at any time compared to 9% to 23% of patients receiving placebo.[83] Dissociation measured by the CADSS generally peaked around 40 minutes following dose administration and resolved by 90 minutes post-dose.[84] Dissociative effects of esketamine were observed to be attenuated with repeat administration.[84]

Sedation

Symptoms related to sedation (e.g., somnolence, sedation) were also commonly reported during the clinical trials of esketamine, occurring more often in patients treated with esketamine than placebo. The Modified Observer's Assessment of Alertness/Sedation (MOAA/S) scale was used to objectively assess sedation during the clinical trials of esketamine. The MOAA/S scale ranges from zero (no response to pain) to five (awake) with scores between zero to four indicating some level of sedation.[84] Across the three short-term trials, MOAA/S scores between zero and four following dose administration were observed in 49% to 61% of esketamine-treated patients at any time compared with 10% to 19% of those treated with placebo.[83] Sedation measured by the MOAA/S scale peaked around 30 to 45 minutes post-dose and generally resolved by 60 to 90 minutes post-dose.[84].

Increased Blood Pressure

In the short-term trials, potential clinically important increases in systolic (≥180 mmHg or increase of ≥20 mmHg) and diastolic (≥105 mmHg or increase of ≥15 mmHg) blood pressure following dose

administration occurred more frequently in patients treated with esketamine compared to those treated with placebo.[84]Across the three trials, 3% and 7% of esketamine-treated patients experienced a potentially clinically important increase in systolic and diastolic blood pressure following dose administration, respectively, compared to 0.3% and 2% of patients treated with placebo.[84] Blood pressure increases peaked at 40 minutes post-dose and generally resolved by four hours post-dose.[83]

Table 3.6. Esketamine: Important TEAEs or TEAE Occurring in ≥ 20% of Patients During the Phase III RCTs

Trial	Arm	N	Any TEAE	SAE	D/C due to TEAE	Nausea	Dissociation	Dizziness	Headache	Vertigo	Somnolence	Sedation	BP Increase
TRANSFORM-1 & -2* [84]	Esketamine	346	87.0	0.9	4.6†	28.3	26.6	23.7	20.2	22.5	17.3	5.5	9.0
	Placebo	222	64.4	0.5	1.4†	8.6	3.6	6.8	17.1	2.3	9.0	0.9	2.3
TRANSFORM-3 [84]	Esketamine	72	70.8	4.2	5.6†	18.1	12.5	22.2	12.5	11.1	NR	NR	12.5
	Placebo	65	60.0	3.1	3.1†	4.6	1.5	7.7	3.1	3.1	NR	NR	4.6

BP: blood pressure, D/C: discontinuation, N: number analyzed, SAE: serious adverse event, TEAE: treatment-emergent adverse event

*Pooled incidence of TEAEs from TRANSFORM-1 & -2 are presented here; †TEAEs leading to d/c of intranasal medication, not antidepressant

Ketamine

We found no trial that directly compared esketamine and ketamine. One Phase II trial found that ketamine provided greater symptom improvement compared to placebo. A greater proportion of patients receiving ketamine also achieved clinical response and clinical remission at two weeks. Important safety events observed were dissociation, dizziness, headache, sedation, and delusion; the FDA label for other indications includes a warning for abuse and dependence.

Clinical Benefits of Ketamine

In the RCT of IV ketamine (Singh 2016), both the twice- and thrice-weekly dosing frequencies of ketamine resulted in a greater reduction in MADRS from baseline to day 15 compared to placebo (twice weekly: mean CFB -18.4 vs. -5.7 [LSMD: -16.0]; thrice weekly: mean CFB -17.7 vs. - 3.1 [LSMD: -16.4]; both p<0.001).[55] Improvement in MADRS from baseline to day 29 was also numerically higher for the twice-weekly and thrice-weekly ketamine groups compared to the corresponding placebo arms, although statistical significance was not reported.[55] Of note, about 80% of participants in the placebo arm had discontinued treatment due to lack of efficacy compared to less than 10% of participants in the ketamine arms by day 29. The considerably high and disproportion rate of discontinuation due to lack of efficacy may reflect a loss of the integrity of the blinding during the trial (see below).

The proportion of participants achieving clinical response at day 15 was higher in both the twice-weekly and thrice-weekly ketamine groups compared to their respective placebo groups (68.8% vs. 15.4%, p=0.005; 53.8% vs. 6.3%, p=0.004, respectively). In addition, numerically more patients in the twice- and thrice-weekly ketamine groups achieved clinical remission compared to their respective placebo groups, but a statistical difference was only observed between the twice-weekly groups (37.5% vs. 7.7%, p=0.05).

We observed that the response and remission rates in the placebo groups of the ketamine trial were much lower compared to the esketamine trials. This could be due to functional unblinding, with subjects realizing they are on placebo. As stated above, there was an unusually disproportionate rate of discontinuation due to lack of efficacy in the placebo groups. This is another reason (in addition to the other study design and population differences described above) we chose not to quantitively compare the esketamine and ketamine trials.

In the single arm study of ketamine designed to assess time to relapse after ketamine discontinuation, 17 of the 24 patients (71%) in the study achieved clinical response (≥ 50% after receiving after receiving six doses of IV ketamine over 12 days.[56] All responders were followed for up to 83 days; the median time to relapse observed was 18 days. Four patients (23.5%) did not relapse by the end of the follow-up phase.

Table 3.7. Ketamine: Symptom Improvement, Clinical Response and Remission at Day 15

Arm	Change in MADRS From Baseline			Clinical Response (≥ 50% Improvement)		Remission (MADRS ≤ 10)	
	Mean Change (SD)	Diff in LSM Change, Mean (SE)	P-Value	Rate, %	P-Value	Rate, %	P-Value
Twice Weekly							
IV Ketamine	-18.4 (12.0)	-16.0 (3.7)	<0.001	68.8	0.005	37.5	0.05
Placebo	-5.7 (10.2)	—	—	15.4	—	7.7	—
Thrice Weekly							
IV Ketamine	-17.7 (7.3)	-16.4 (2.4)	<0.001	53.8	0.004	23.1	0.08
Placebo	-3.1 (5.7)	—	—	6.3	—	0	—

Diff: difference, IV: intravenous, LSM: least square mean, MADRS: Montgomery-Åsberg Depression Rating Scale, SD: standard deviation, SE: standard error

Harms of Ketamine

In Singh 2016, a larger proportion of participants receiving ketamine experienced any AEs and drug-related AEs compared to those receiving placebo (Table 3.8).[55] Similar to the esketamine trials, the most common AEs reported during the double-blind phase included nausea, dissociation, dizziness, and anxiety (Table 3.8). Dissociative symptoms as assessed by CADSS were noted to peak at 40 minutes after the start of infusion and resolved by 3 hours post infusion. And as noted in the esketamine trials, the intensity of dissociative symptoms was reduced with repeated dosing of ketamine. Ketamine was also observed to be associated with increased psychotomimetic symptoms (delusion or delirium), as assessed by Brief Psychiatric Rating Scale positive symptom subscale (BPRS+). This also generally returned to pre-infusion values about 3 hours following infusion.

Two participants (11.1%) receiving ketamine experienced SAEs (anxiety and suicide attempt) compared to no SAEs reported for participants receiving placebo. Neither event was determined to be related to the study drug. No deaths were reported.

A similar pattern of side effects was observed in the single-arm, long-term trial, with patients experiencing dissociative and psychotomimetic symptoms that generally resolved four-hours post-infusion.[56] The most commonly reported side effects during the four-hour post-infusion period included feeling strange or unreal (58.3%), abnormal sensation (54.2%), blurred vision (50.0%), and drowsy or sleepiness (45.8%).

Although no evidence of misuse or abuse was reported in the TRD trials of ketamine, ketamine has been reported as a drug of abuse due to its dissociative and hallucinogenic effects.[87] The current FDA label for ketamine (for other indications) includes a warning for drug abuse and dependence.[88]

Table 3.8 Proportion of Patients Experiencing Adverse Events in RCT of Ketamine

Arm	N	Any AE	Related AE	SAE	D/C due to AE	Nausea	Dissociation	Dizziness	Headache	Anxiety
Ketamine twice weekly	18	83.3	72.2	11.1	11.1	16.7	27.8	22.2	22.2	27.8
Placebo twice weekly	16	56.3	37.5	0	6.3	6.3	0	6.3	31.3	0
Ketamine thrice weekly	17	76.5	58.8	0	5.9	23.5	5.9	11.8	41.2	5.9
Placebo thrice weekly	16	50.0	31.0	0	0	12.5	0	0	6.3	0

AE: adverse event; D/C: discontinuation; N: number analyzed; SAE: serious adverse event

Other Comparators: rTMS, ECT and Augmentation with Olanzapine

We found no trials that compared esketamine to rTMS, ECT or augmentation with olanzapine.

In the 11 sham-controlled trials of rTMS, rTMS resulted in numerically greater improvement from baseline on MADRS and/or HAM-D score compared to sham at four to six weeks, however statistically significant differences were observed in only five of the trials. Two of the trials found no difference between rTMS and sham, while four studies did not report statistical significance. Similar trends were observed for remission and clinical response outcomes. Among the trials that used the MADRS scale, the difference in symptom improvement observed between rTMS treated patients compared to sham treated patients ranged from a score of -2 to -7 at four to six weeks. The most commonly reported AEs with greater occurrence in the rTMS treated patients were scalp discomfort, pain and headache.

In the small RCT that compared ECT with rTMS (42 patients), no difference was observed in the efficacy of both interventions based on symptom improvement, clinical response (40% vs. 50% respectively), and remission rates (20% vs. 10%) at four weeks.[82] Although not reported in the trial of ECT included in our review, the FDA label of ECT includes a warning for disorientation, confusion, memory problems, pain, skin burns, physical trauma, seizures, pulmonary complication, cardiovascular complications and death.[89]

In the two similarly designed studies of olanzapine, there was no difference observed in symptom improvement, remission rates and clinical response rates between olanzapine/fluoxetine treated patients and placebo plus antidepressant treated patients at eight to 12 weeks.[68,69] Patients in the olanzapine/fluoxetine arm observed a higher incidence of somnolence, peripheral edema, weight gain and increased appetite compared to patients randomized to placebo plus antidepressants arms. Discontinuation due to weight gain occurred at a higher incidence in olanzapine/fluoxetine treated patients compared to all other groups.

See Appendix Table D7 for additional details on each study.

Controversies and Uncertainties

Several important limitations in the available evidence about the comparative benefits and harms of esketamine for patients with treatment-resistant depression are worth highlighting. Though many studies include patients having failed two or more therapies in the current episode, this definition is not uniform, and as a result we found heterogeneity in the studies we reviewed in terms of the severity of the MDD episode. This definition of TRD also only applies to the current episode and does not consider the number of past episodes, their severity or duration. This is important when considering which patients with TRD may most benefit from esketamine. Clinical experts we spoke with viewed that esketamine may be an option for patients with chronic, severe depression who have failed multiple other therapies.

We identified three, phase III randomized placebo-controlled trials of esketamine for short-term use in patients with TRD (TRANSFORM-1, -2, & -3). Patients were required to have failed two therapies in the current episode including one that could have been given during a four-week prospective screening and observational phase. How patients included in this study reflect the very severe patients that experts felt would be the ones they would consider for esketamine is unclear since only 36-40% had been on and failed 3 or more medicines during the current episode. Each of these trials compared esketamine to placebo along with the addition of a new antidepressant (an SSRI or SNRI) at the clinician's discretion. Thus, these trials compare what may be considered the additive benefit and harm of esketamine rather than directly comparing esketamine to the use of an antidepressant. Moreover, we could find no studies directly comparing esketamine to other therapies used in patients with TRD including augmentation with medications such as antipsychotics, as well as TMS and ECT.

Patients with MDD may have other co-existing psychiatric illnesses such as bipolar disorder (termed depression with mixed features), substance use disorders and anxiety disorders. Patients with TRD who have such other psychiatric conditions may not respond as well to antidepressant treatments. It is unclear how esketamine may work in such patients. Patients with co-existing disorders including psychosis, mania, and moderate or severe substance use were excluded. Subgroup

analyses of available data in the esketamine trials have not yet been published describing patient outcomes among those who had other psychiatric conditions, such as anxiety disorders.

We sought to indirectly assess the comparative benefits of esketamine to other therapies using network meta-analysis. Specifically, we sought to compare esketamine with ketamine, other antidepressants, augmentation medications, TMS and ECT. Though we found trials for some of these comparators in patients with TRD, differences in key aspects of these trials precluded our ability to perform a network meta-analysis. These included important differences in entry criteria, study populations, study design and outcome measurements across these trials. As a result, we did not think it appropriate to perform a network meta-analysis across the trials. Instead, we compared the benefits and harms of esketamine to placebo plus background (either new or continued) antidepressants.

The three, phase III randomized placebo-controlled trials of esketamine for short-term use in patients with TRD (TRANSFORM-1, -2, & -3) all reported improved outcomes among patients randomized to esketamine, but in only one trial (TRANSFORM-2) was the primary outcome comparison statistically significant. This may cause uncertainty about the benefits of esketamine. Since the TRANSFORM-3 trial involved a different study population, patients 65 years of age and older, we conducted a meta-analysis of data from two of the esketamine trials (TRANSFORM-1 & -2) that were homogenous in terms of inclusion and exclusion criteria, study design, and outcome. The key difference between these two trials was that TRANSFORM-1 involved a fixed dose schedule comparing 56 mg and 84 mg of esketamine versus placebo, while TRANSFORM-2 permitted flexible dosing starting with 56 mg and increasing to 84 mg based upon patient response. Our meta-analysis demonstrated that esketamine plus antidepressant resulted in greater symptoms improvement and more patients achieved a clinical response and a clinical remission than placebo, but statistical significance was not reached for clinical remission. These results support our conclusion that esketamine provides a short-term benefit in patients with TRD.

Given the chemical similarity between ketamine and esketamine, we were interested in comparing the clinical benefit and harms of ketamine in the available trials. Ketamine is primarily given by IV infusion, but patients and experts describe the use of intranasal ketamine as well. However, we were not able to find any trials of intranasal ketamine that met our eligibility criteria. One relevant trial of ketamine was identified, but differences in the placebo response rate led to us not performing a network meta-analysis with esketamine . It is important to note that while the outcomes reported in the groups treated with esketamine and ketamine were of similar magnitude, the placebo response rate was much higher in the esketamine than the ketamine trial. Performing a network meta-analysis would have led to concluding that ketamine was significantly more effective than esketamine, mainly driven by the lower placebo response rate. Though we did not perform a network meta-analysis, we did develop a cost-utility model evaluating esketamine and ketamine given the similar treatment outcomes and the widespread use of off-label ketamine infusion clinics for patient with TRD.

There is also uncertainty about the long-term use of esketamine for patients with TRD. The SUSTAIN-1 trial examined relapse in patients who reported an initial response to esketamine. The study showed higher rates of relapse among patients who discontinued esketamine compared to those who continued to take it. These outcomes support the need for long-term therapy and are also reflected in what we heard from patients and experts. Specifically, patients with TRD who respond to a new therapy are likely to be continued on it for a prolonged period of time. This reflects the long duration of depression symptoms and the lack of response to prior therapy or side effects limiting the use of such therapies. The SUSTAIN-2 trial examined the open label use of esketamine for up to 48 weeks. Side effects and discontinuation rates were low, which is reassuring. However, the long-term comparative benefits of esketamine are unknown.

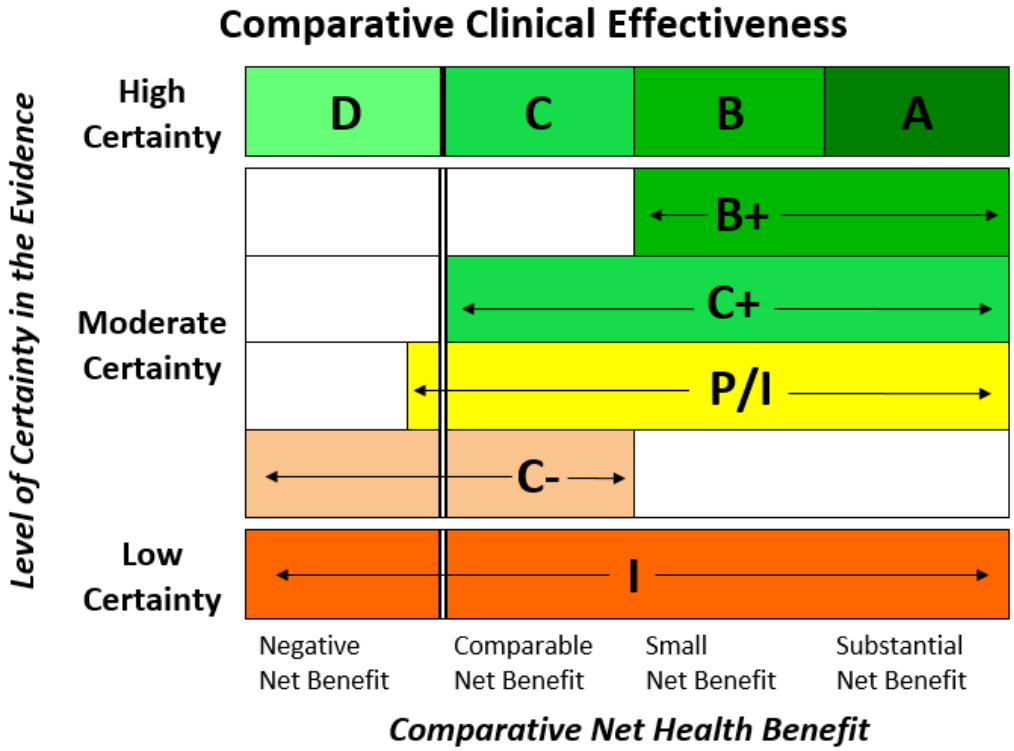
Though the esketamine trials did not report issues related to misuse or abuse, this remains a concern given the similarity to ketamine, which is reported to have these risks. For this reason, esketamine is classified also a Schedule III substance.[84] It is unclear from available information how misuse and abuse were evaluated in the esketamine trials. Despite the lack of concern from trial data, esketamine will be made available only through a Risk Evaluation and Mitigation Strategies (REMS) program in order to monitor its abuse potential. Thus, its long-term safety continues to include concerns about its potential for misuse or abuse.

While esketamine when combined with a new antidepressant appears to offer favorable short-term results compared to placebo plus a new antidepressant, the long-term benefits and harms remain unclear. Since most patients with TRD will require maintenance therapy to control their symptoms, it remains to be seen how esketamine will be used in routine practice. The SUSTAIN-1 trial demonstrates that relapse will be common if esketamine is discontinued. Thus, short-term control of symptoms with use of esketamine means that if it is stopped, other therapies will need to be added. Given that many patients have already failed these other therapies, clinicians will likely need to use esketamine for maintenance therapy despite the lack of data, especially compared to alternative treatments.

Finally, given the impact of TRD on quality of life, patient reported measures were included as secondary outcomes in the esketamine trials. These outcomes including quality of life also demonstrated improvement with esketamine. Patients and patient advocates have highlighted the importance of TRD on quality of life and measures of work and productivity and challenges in adequately measuring their impact. Given these limitations in measuring the quality of life in patients with TRD and how it may change with treatment, there is uncertainty regarding the magnitude of benefit for treatments of TRD on patients' overall quality of life. Developing validated, sensitive measures that can sufficiently capture the individual burden of the disorder in all affected patients remains an important challenge.

3.4 Summary and Comment

Figure 3.4. ICER Evidence Rating Matrix



A = "Superior" - High certainty of a substantial (moderate-large) net health benefit
B = "Incremental" - High certainty of a small net health benefit
C = "Comparable"- High certainty of a comparable net health benefit
D = "Negative"- High certainty of an inferior net health benefit
B+ = "Incremental or Better" - Moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit
C+ = "Comparable or Better" - Moderate certainty of a comparable, small, or substantial net health benefit, with high certainty of at least a comparable net health benefit
P/I = "Promising but Inconclusive" - Moderate certainty of a comparable, small, or substantial net health benefit, and a small (but nonzero) likelihood of a negative net health benefit
C- = "Comparable or Inferior" - Moderate certainty that the point estimate for comparative net health benefit is either comparable or inferior
I = "Insufficient" - Any situation in which the level of certainty in the evidence is low

Table 3.9. ICER Ratings on the Comparative Net Health Benefit of Esketamine*

Interventions	Background Antidepressant Alone	Ketamine	ECT, TMS	Augmentation with Olanzapine
Esketamine Plus Background Antidepressant	P/I	I	I	I

Esketamine Versus Placebo Plus Background Antidepressants

- In adults (ages 18 to 64 years) on newly initiated background antidepressant, symptom improvement at four weeks was greater with esketamine than placebo. More patients also achieved clinical response and clinical remission on esketamine compared to placebo; however statistical significance was not reached for clinical remission.

- In adults ages 65 and older on newly initiated background antidepressant, symptom improvement at four weeks was not significantly different between esketamine and placebo; however, the magnitude of improvement observed with esketamine in this population was comparable to what was observed in adults ages 18 to 64 years.

- In adults (ages 18 to 64 years) who achieved stable clinical remission or stable clinical response (without remission), continued treatment with esketamine plus background antidepressant, as assessed over a 48 weeks period reduced the risk of relapse compared to switching to placebo plus background antidepressant.

- Esketamine was generally well tolerated in the short-term Phase III trials, however, there were important safety concerns such as dissociation and increased blood pressure associated with esketamine use along with risk of suicide. In addition, although there was no evidence of abuse and misuse during the trials, these remain an important safety concern, due to esketamine's pharmacological similarity to ketamine, a drug that has been reported to be abused and misused for its dissociative and hallucinogenic effects. There is limited data on long-term use of esketamine.

In summary, the results of the Phase III trials show that esketamine is promising in terms of clinical efficacy for symptom improvement and achieving clinical response compared to placebo. However, in the absence of long-term safety data, we cannot definitively rule out the possibility of a small net harm. Thus, for adults (18 years and older) with TRD, we consider the evidence on esketamine plus background antidepressant compared to background antidepressant alone to be "promising but inconclusive" (P/I), demonstrating a moderate certainty of a comparable or substantial net health benefit, and a small (but non-zero) likelihood of a negative net health benefit.

Esketamine Versus Ketamine, TMS, ECT and Augmentation with Olanzapine

We attempted to compare esketamine with ketamine, ECT, TMS, oral antidepressants, or augmentation with antipsychotics (e.g., olanzapine). However, we did not identify any head-to-head evidence comparing esketamine with any of these comparators. In addition to a lack of comparative data, differences in entry criteria, patient characteristics, study design and outcome measurement in the clinical trials of esketamine and these comparators precluded even indirect comparison through network meta-analysis. Thus, we feel the evidence is insufficient ("I") to judge

the net health benefit of esketamine versus ketamine, ECT, TMS, oral antidepressants, or augmentation with antipsychotics (e.g., olanzapine).

4. Long-Term Cost Effectiveness

4.1 Overview

The primary aim of this economic evaluation was to estimate the cost-effectiveness of the addition of esketamine nasal spray to a new oral antidepressant compared to no additional treatment (new oral antidepressant alone), in patients receiving a newly prescribed oral antidepressant, for the treatment of treatment-resistant major depressive disorder (TRD) using a *de novo* decision analytic model. These two strategies in the analysis are referred throughout as "esketamine" and "no additional treatment." Both strategies include subsequent lines of antidepressant therapy following discontinuation of primary intervention(s). The outcomes of interest included the incremental cost per quality-adjusted life year (QALY) gained, life-year (LY) gained, and depression-free day. All costs and outcomes were discounted at a rate of 3% per year. For this aim, the base-case analysis was conducted using a health care sector perspective (i.e., focus on direct medical care costs only) and a lifetime horizon when evaluating cost per QALY and cost per LY gained. For the cost per depression-free day outcome, only the direct treatment effects (i.e., those patients who obtained remission on either the esketamine treatment pathway or oral antidepressant treatment pathway were considered. For this cost per consequence analysis, a shorter two-year time horizon was employed, because of a high esketamine discontinuation rate and uncertainty over long-term use. Productivity gains with effective treatment were considered in a separate scenario analysis. The model was developed in Microsoft Excel 2016 (Redmond, WA).

A review of the literature of potential comparators to esketamine was conducted. When the available trials were reviewed, one trial comparing ketamine to placebo, with patients continuing any other antidepressant medications they were receiving at study entry, emerged for potential inclusion in a network meta-analysis (NMA). However, further evaluation of the trial revealed that while ketamine had similar rates of response and remission to studies evaluating esketamine, placebo response and remission were very different from esketamine placebo trials. These differences suggest that either the enrolled patient populations differed greatly, treatments were administered or evaluated in very different ways, and/or other factors affecting the placebo response may be present. The heterogeneity present in these trials was deemed too substantial to conduct an NMA from which to derive needed inputs for a cost-effectiveness model. However, given the similar, but non-comparable efficacy of ketamine and esketamine, a cost-analysis was undertaken to provide payers and others with some estimate of differences in expected costs for each of these treatments.

Thus, as a secondary aim of this report, we evaluated the one-year costs of treatment with esketamine compared to intravenous ketamine for the treatment of TRD. For this aim, the base-case analysis was conducted using a health care sector perspective, focusing on direct medical care

and patient out-of-pocket costs, with a one-year time-horizon. No discounting was applied to this cost analysis. Productivity gains with effective treatment were considered in a scenario analysis. The model was developed in Microsoft Excel 2016 (Redmond, WA).

4.2 Methods

Model Structure

For the cost-effectiveness analysis, we developed a *de novo* decision analytic model informed by key clinical trials and prior relevant economic models. The base-case analysis was from the perspective of the health care sector and thus focused on direct medical care costs only. Costs and outcomes were discounted at 3% per year.

The model focused on an intention-to-treat analysis, with a hypothetical cohort of patients with TRD, all being treated with either esketamine or no additional treatment upon entry into the model. The model cycle length was set at three months, based on the length of typical treatment initiation seen in trials for treatment-resistant depression.

Definitions of different levels of treatment effectiveness are available in Table 4.1. As shown in the model schematic, Figure 4.1, simulated patients entered the model with severe depression receiving an "Initial Treatment" (i.e., antidepressant + esketamine or antidepressant + no additional treatment). Initial treatment may have resulted in a considerable improvement in depression symptoms (Markov state "Initial treatment effective, remain on initial treatment, no depression"), a lesser improvement in depression symptoms (Markov state "Initial treatment partly effective, remain on initial treatment with augmentation, mild to moderate depression"), or an insufficient response (Markov states "Initial treatment not effective, switch to alternative treatment 1 (effective), no depression" or "Initial treatment not effective, switch to alternative treatment 1 (not effective), severe depression").

Table 4.1. Treatment Response Definitions Used in the Model

Model State Description	Definition	Calculation from Clinical Trials
Treatment Effective	MADRS score of 12 or less or QIDS-SR_{16} of 5 or less	Proportion achieving remission
Treatment Partly Effective (applies only to initial treatments)	50% reduction from baseline MADRS score, but not achieving a MADRS score of 12 or less or 50% reduction from baseline QIDS-SR_{16}, but not achieving a QIDS-SR_{16} of 5 or less	Proportion achieving response – Proportion achieving remission
Treatment Not Effective or Treatment Loses Effect	Less than a 50% reduction in MADRS or QIDS-SR_{16}, score when compared with baseline	Those not achieving response (i.e., 1 – response)

Each cycle, patients whose initial treatment was effective may continue to experience effective treatment or may lose initial treatment effect. Those with continued treatment effectiveness will typically remain in the same Markov state ("Initial treatment effective, remain on initial treatment, no depression"), although some patients with continued response may have the initial treatment discontinued and remain without depression. Those patients who responded to the initial treatment, but subsequently had a loss of effect will transition to one of two Markov states, depending on the effectiveness of the subsequent alternative treatment (Markov states "Initial treatment not effective, alternative treatment 1 (effective), no depression" or "Initial treatment not effective, alternative treatment 1 (not effective), severe depression").

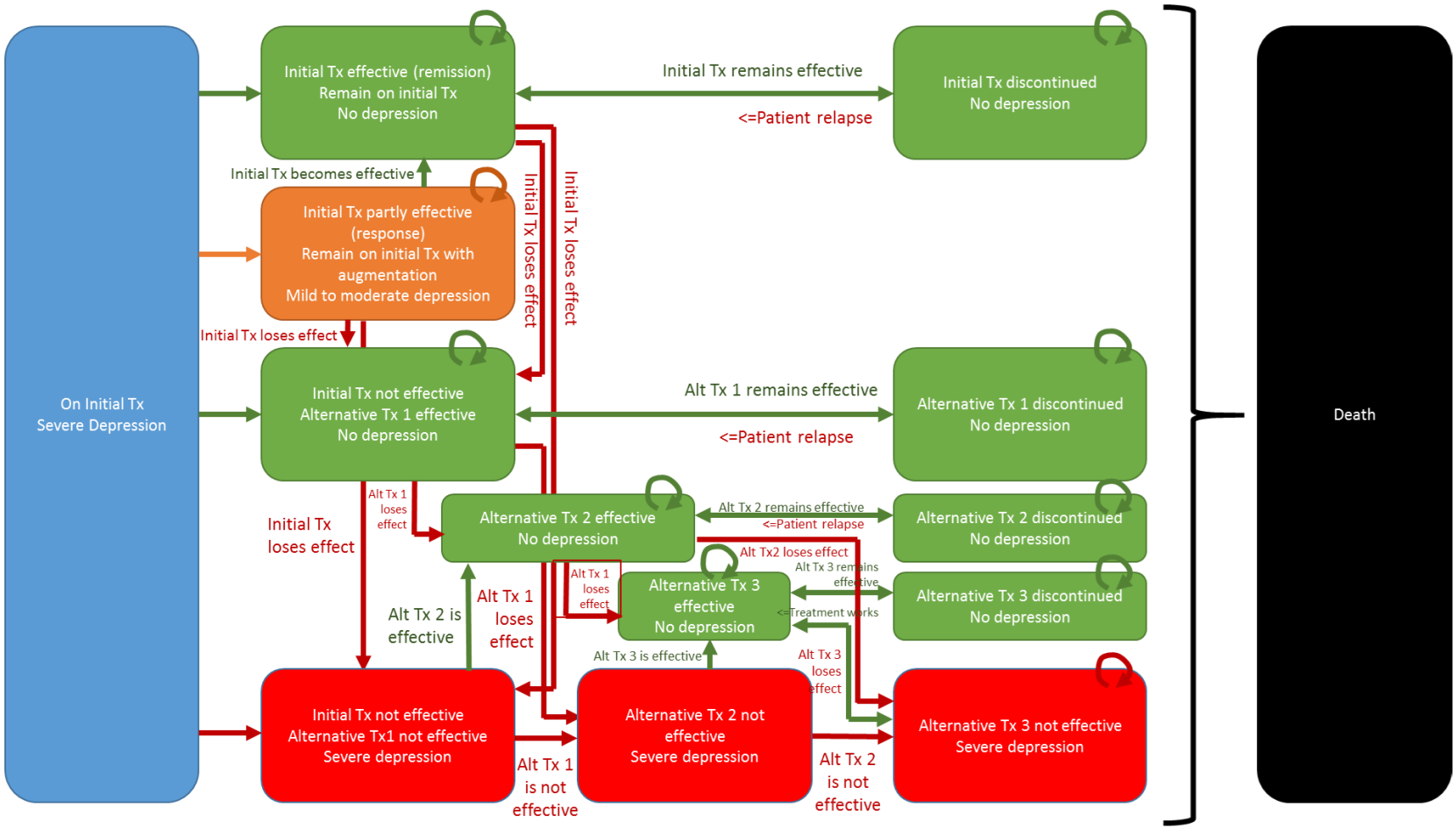
Patients whose initial treatment was partly effective received augmentation added to their initial treatment. Each cycle, patients whose initial treatment was partly effective may continue treatment with initial treatment plus augmentation (Markov state "Initial treatment partly effective, remain on initial treatment with augmentation, mild to moderate depression"), may have initial treatment become more effective (Markov state "Initial treatment effective, remained on initial treatment, no depression"), or subsequently have a loss of treatment effect thereby transitioning to one of two Markov states depending on the effectiveness of the subsequent alternative treatment (Markov states "Initial treatment not effective, alternative treatment 1 (effective), no depression" or "Initial treatment not effective, alternative treatment 1 (not effective), severe depression"). Patients in whom the initial treatment was partly effective long-term were not allowed to have their initial treatment discontinued while in this partial response Markov state.

Patients not responding to the initial treatment transitioned from the initial treatment state to one of two Markov states, depending on the effectiveness of the alternative treatment. Those that responded to the first alternative treatment were in the "Initial treatment not effective, alternative treatment 1 (effective), no depression" Markov state in the second model cycle. Those that did not respond to the alternative treatment were in the "Initial treatment not effective, alternative treatment 1 (not effective), severe depression" Markov state. In subsequent cycles, patients may have transitioned to up to three alternative treatments if they experienced a loss of treatment effect with current treatment. As with effective initial treatment, most patients with effective treatment over the longer term continued to take the alternative treatment and remained in the Markov state "Initial treatment effective, switch to alternative treatment 1-3 (effective), no depression" corresponding to which alternative treatment was effective. A small number of patients who experienced continued effect to the respective alternative treatment may have moved to a state where the alternative treatment was discontinued with no depression (Markov states "Alternative treatment 1-3 discontinued, no depression."

Any patient with continued response to initial treatment or any alternative treatment, and for whom treatment was discontinued, there was a possibility of relapse into depression. These patients transitioned back to their most recent effective treatment.

Patients remained in the model until they died. All patients transitioned to death from all causes from any of the alive health states (please see Mortality section below for further clarification).

Figure 4.1. Model Framework



Blue = initial treatment; Green = treatment effective; Orange = treatment partly effective; Red = treatment not effective

Note: Double sided green arrows demonstrate two transitions, patients who discontinue therapy because of long-term treatment effect and patients who have a remission and need to restart treatment with the last effective therapy.

Target Population

The population of focus for the economic evaluation were adults with a mean age of 46 years. Patients entered the model with either a single episode or recurrent major depressive disorder without psychotic features that was treatment-resistant. Treatment-resistance was defined as non-response to two or more adequate trials of antidepressant treatment in the current depressive episode.[5] Baseline patient characteristics are presented in Table 4.2.

Table 4.2. Base-Case Model Cohort Characteristics

Baseline Characteristics	Value	Source
Mean Age, Years (SD)	46 years	TRANSFORM-1 & -2[50,51]
Female, %	67%	TRANSFORM-1 & -2[50,51]
Number of Previous Antidepressant Trials, % 1 or 2 ≥3	 63% 37%	TRANSFORM-1 & -2[50,51]
MADRS Score at Baseline, Mean	37.4	TRANSFORM-1 & -2[50,51]

Treatment Strategies

The modeled treatment strategies were based on trial data.[50,51] The intervention included was esketamine (Spravato™, Janssen) 56 mg or 84 mg administered intranasally twice weekly, reduced to once weekly or every other week, plus an unspecified new oral antidepressant agent. This treatment arm with esketamine included subsequent lines of therapy with oral antidepressants following esketamine's discontinuation. The comparator was intranasal placebo administered intranasally twice weekly plus an unspecified new oral antidepressant agent. In the model, this represented a treatment pathway comprising of multiple lines of treatment with oral antidepressants, without esketamine.

Based on the judgement of clinical experts, esketamine treatment was viewed as an option for later-line treatment after patients had failed numerous oral antidepressants. Potential comparators included electroconvulsive therapy, transcranial magnetic stimulation, and ketamine. However, systematic differences in study design, heterogeneity between patient populations, and inconsistency in the outcomes assessed by clinical trials evaluating these therapies precluded the inclusion of these comparators in an NMA. As a result, it was not possible to generate effect estimates for other therapies compared to esketamine. Therefore, the model compared esketamine plus a new oral antidepressant to a new oral antidepressant alone (i.e., the placebo comparison arms of the TRANSFORM-1 & -2 studies).[50,51]

Key Model Characteristics and Assumptions

The model required several assumptions. Key model assumptions and rationale for the assumptions are presented in table 4.3.

Table 4.3. Key Model Assumptions

Assumption	Rationale
Some patients with effective treatment in each three-month cycle had their treatments discontinued.	In the treatment of major depressive disorder, patients with sustained response to treatment often have their treatments discontinued. There is limited information regarding the frequency of this practice in patients with treatment-resistant depression. We assumed that some patients had their treatments discontinued, with the probability based on expert opinion. This input was subjected to a robust sensitivity analysis.
Patients who had discontinued treatment due to sustained effectiveness, but then subsequently relapsed, restarted their last effective treatment and were assumed to receive immediate benefit from that treatment.	There is limited information regarding practice treatment patterns in patients with treatment-resistant depression. Restarting patients on therapies that were previously effective is a common practice in major depressive disorder. We assumed immediate benefit to the treatment to keep the model simple. This assumption affects a small proportion of modeled population and is unlikely to have a measurable effect on the model estimates.
Patients in whom initial treatment was only partly effective had mild to moderate depression.	In the TRANSFORM-1 & -2 trials, the minimum starting MADRS score was 28 or greater. Response was defined as at least a 50% reduction in the MADRS score without achieving remission (defined as a MADRS score of less than 12).[50,51] Given a maximum MADRS score of 60 and a mean MADRS score of 37 - 38, patients with response without achieving remission would have scores between 12 and 30, which correspond to mild (scores 9-17) or moderate (scores 18-34).[50,51]
Patients in whom initial treatment was only partly effective continued treatment with their initial treatment and received augmentation.	There is limited information regarding practice treatment patterns in patients with treatment-resistant depression. The STAR*D study allowed patients to receive augmentation or switch antidepressants for patients in whom treatment was only partly effective and depression was still present.[12] As we did not have probabilities for the proportion of patients likely to choose augmentation vs. switch treatment and since esketamine was generally well tolerated in its key trials, we assumed that patients would remain on partly effective treatments with augmentation.[50,51,90]

Assumption	Rationale
Patients in whom treatment was not effective discontinued treatment and received an alternative treatment.	There is limited evidence regarding treatment patterns in patients with treatment-resistant depression, especially as it relates to patients who receive some benefit but experience a suboptimal response to a new treatment. Clinical trials, including TRANSFORM-1 & -2 and STAR*D, considered non-response as those patients who did not achieve "remission" nor "response." In the STAR*D, some patients who achieved response switched treatment.[12] To simplify the model, we assumed that only those patients who did not achieve remission or response, either due to lack of treatment response or discontinuation of treatment, received a different treatment.
The Markov state "Alternative treatment 3 not effective, severe depression" represented the third and all future treatments that were not effective. Simulated patients remained in that state long-term if all future therapies were not effective or moved to the Markov state "Alternative treatment 3 effective, no depression" if a future alternative treatment was effective.	Costs were not evaluated for patients requiring eight or more regimen changes. Probabilities for treatment failure were not available beyond four therapies. Since patients were entering the model already having failed an average of three therapies (with the new treatment being at least the third treatment), failing more than three additional alternative therapies resulted in the same costs for each additional failed treatment. Therefore, we decided to limit the number of alternative treatments in the model to three.
Treatment does not directly affect mortality.	The TRANSFORM-1 & -2 trials did not evaluate the impact of esketamine on mortality. However, depression has been linked with a higher mortality rate. We adjusted all-cause mortality for those with treated vs. untreated depression.
Modeled costs were associated with number of previous therapies and not directly with depression severity.	Cost data was not available evaluating the total costs of treating treatment-resistant depression by disease severity. Cost data was available according to the number of failed therapies. The model was developed to incorporate data that was available from the literature.
Patients with effective depression treatment had medical costs (not including pharmaceutical costs) equivalent to those with three prior treatment failures (i.e., on their fourth treatment).	Comparative cost data for patients with and without adequately treated treatment-resistant depression was not available. We therefore assumed that the lowest available cost from Russell et al. should apply to all patients on effective treatment with initial treatment.[91]
Patients who were effectively treated for at least one cycle (i.e., three months) and had their effective treatment discontinued incurred prescription costs equivalent to those with three prior treatment failures (i.e., on their fourth treatment).	Comparative cost data was not available for patients with and without adequately treated treatment-resistant depression who discontinued therapy. We therefore assumed that the lowest available prescription cost from Russell et al. applied to all patients who were effectively treated and had their most recent treatment discontinued.[91]

Model Inputs

Clinical Inputs

Short-term clinical inputs of the relative risk of depression remission and response for esketamine and the comparator were derived from a meta-analysis of the esketamine clinical trials TRANSFORM-1 & -2.[50,51] Long-term clinical inputs related to continued response of esketamine were derived from the SUSTAIN-1 study.[53] Long-term clinical inputs related to alternative oral antidepressant treatments were derived from the STAR*D trial.[12]

Clinical Probabilities/Response to Treatment

The decision model was evaluated over a lifetime time horizon with 3-month cycles. Patients began with severe depression and received initial treatment with a new oral antidepressant plus esketamine or new oral antidepressant with no additional treatment. The degree of response to esketamine or no additional treatment (i.e., placebo) was based on clinical trials evaluating outcomes at four weeks, in which remission was defined as achieving a MADRS score of 12 or less at four weeks and response was defined as achieving a 50% or greater reduction in the MADRS score from baseline at four weeks. A selected list of inputs is shown in Tables 4.4 and 4.5 below. Since esketamine's treatment effect was similar at week four of the TRANSFORM-1[50] trial and three months after initiation (i.e., week four of the SUSTAIN-2 trial),[54] four-week estimates for effective and partly effective treatment probabilities were used to represent three-month transition probabilities and were not transformed. The probability of non-response to esketamine and subsequent effective or ineffective treatment with an alternative treatment was calculated from the weighted probability of non-response to esketamine from the TRANSFORM-1 & -2 trials[50,51] and the probability of achieving remission with an alternative treatment at the next treatment step, derived from the STAR*D trial.[12] The probability of discontinuing esketamine or the comparator after long-term treatment success was obtained from expert opinion.

For those in whom esketamine or the comparator was effective long-term and discontinued, the probability of losing effectiveness was based on estimates from the STAR*D trial.[12] Patients received the last effective treatment (esketamine or the comparator), which was assumed to be effective in treating the recurrence of depression.

For patients in whom esketamine or the comparator was effective or partly effective and who were continuing treatment, estimates of loss of effect were obtained from the SUSTAIN-1 trial.[53] For those patients in whom esketamine or the comparator was partly effective, results from the SUSTAIN-1 trial was used to estimate the probability of effective treatment (regression output from SUSTAIN-1 trial, results included in open input document) or loss of effect (stable responders who experienced relapse at 12 weeks).[53] For those who lost effect, these probabilities were multiplied by

the probability of effective treatment with an alternative treatment at the next treatment step, derived from the STAR*D trial.[12] The full calculation is shown in the Appendix Table E2.

Where inputs are available only from Kaplan Meier (KM) curves or bar graphs, probabilities were derived using a digitized estimate value on the KM curve at the appropriate time point. Where probability estimates were not available at three months (i.e., the model's cycle length), probabilities were transformed to three-month probabilities using the appropriate form of the equation $P[t]=1-e^{-rt}$, where P[t] is the probability at time t, r is the corresponding constant rate, and t is the time period over which the probability was assessed. Exceptions to transforming probabilities to three-month probabilities were made for initial response to esketamine (as described above) and the probability of achieving remission on alternative therapy, for which necessary data was not available (i.e., the timing at which the 13.0% of Step 4 patients achieved remission) from the STAR*D trial table 4.[12]

Table 4.4. Treatment Dependent Three-Month Transition Probabilities Used in the Model Derived from Meta-Analysis

Model Input	Esketamine	No Additional Treatment	Source
Remission, Relative Ratio (95% CI)	1.37 (0.99-1.91)	Comparator	Meta-analysis of TRANSFORM-1 & -2[50,51]
Effective Initial Treatment, Probability (95% CI)	39.5% (28.5% – 55.0%)	28.8%	Meta-analysis of TRANSFORM-1 & -2, calculated from RR[50,51]
Partly Effective Treatment, Relative Ratio (95% CI)	1.30 (1.08-1.56)	Comparator	Meta-analysis of TRANSFORM-1 & -2[50,51]
Partly Effective Initial Treatment, Probability (95% CI)	19.3% (9.5% – 31.2%)	16.5%	Meta-analysis of TRANSFORM-1 & -2, calculated from RR[50,51]
Probability of Patients with Initial Partial Response Achieving Complete Response	19.9%	12.4%	SUSTAIN-1 (calculated from long-term relapse and remission rates)[53]
Probability of Patients with Initial Partial Response Losing Response	21.0%	47.6%	SUSTAIN-1 (calculated from long-term relapse and remission rates)[53]

Table 4.5. Non-Treatment Dependent Three-Month Transition Probabilities Used in the Model

Model Input	Value	Source
Probability of Loss of Initial Treatment Effectiveness	13.0%	SUSTAIN-1[53]
Probability of Effective Treatment with Alternative Treatment	13.0%	STAR*D (step 4 from table 3)[12]
Proportion of Patients with Long-Term Effectiveness Discontinuing Treatment	1.3% per cycle (5% per year)	Expert opinion
Proportion of Patients Dying	Age-specific, adjusted for depression	USA Human Mortality Database[92] Ruetfors 2018.[93]

Discontinuation

Discontinuation of esketamine due to treatment-emergent adverse events occurred in 9.5% of patients receiving esketamine and 4.1% of patients receiving antidepressants in the SUSTAIN-2 open label trial.[54] Discontinuation of treatment with alternative oral antidepressants varies by specific agent used. Discontinuation of esketamine, the comparator, or alternative treatments was assumed to be embedded in loss of treatment effect from clinical trials. Therefore, treatment discontinuation specifically due to treatment-emergent adverse events was not explicitly incorporated into the model but was implicitly captured through treatment changes due to loss of treatment effect. Discontinuation of effective treatment was assumed to be 5% per year (1.3% per 3 months) based on clinical expert opinion.

Mortality

Table 4.6 shows mortality inputs used in the model. Gender and age-specific all-cause mortality was sourced from the US tables of the Human Mortality Database.[92] Mortality rates were adjusted to reflect increased all-cause mortality for patients with untreated treatment-resistant depression, smoothed using a moving average approach.[93]

Table 4.6. Mortality Inputs

Parameter	Value	Source
Annual All-Cause Mortality	Varies by age and gender	
Male, 46 Years Old (33% of patients)	0.35%	USA Human Mortality Database[92]
Female, 46 Years Old (67% of patients)	0.22%	
Weighted Average, 46 Years Old	0.27%	
Adjusted Excess Mortality Rate Ratios for Patients with Treatment-Resistant Depression		Ruetfors 2018[93]
Age		
18-29 Years	2.20	
30-49 Years	1.62	
50-69 Years	1.25	

Health State Utilities

Table 4.7 shows health state utilities used in the model. Utilities were derived from two sources; both of which used the Euroqol 5-D questionnaire (EQ-5D). Utility for patients with effectively treated depression were derived from the US population average utility, weighted by gender for our modeled population.[94] The population evaluated in the study used to estimate utility for patients with mild to severe depression were derived from baseline data consisting of individuals with major depressive disorder enrolled in a study evaluating transcranial magnetic stimulation.[95] Characteristics of patients in this study were similar to those in the TRANSFORM trials, with an average number of 3.6 treatment failures and PHQ-9 score of 18.3. Patient EQ-5D health index scores were measured at baseline and stratified according to baseline PHQ-9 levels of mild (<10), moderate (11 to 15), moderately-severe (16 to 20), and severe depression (>20). For the purposes of our study, PHQ-9 severity level was dichotomized into mild to moderate depression and moderately-severe to severe depression, then converted to the equivalent MADRS severity level stimulation using a crosswalk provided in the open input period (open input from Janssen).[96] After applying the crosswalk between the PHQ-9 and MADRS scales, the cutoffs describing depression severity from Janicak et al. were not identical to those from the TRANSFORM-1 & -2 trials. We therefore calculated a weighted utility for patients with mild to moderate depression and moderately severe to severe depression.

The onset of benefit with esketamine occurs rapidly, within approximately one week of initiating treatment. Response to the placebo was also observed quickly, but with a lesser impact on the mean MADRS score at each time point when compared with esketamine. The area between the esketamine and placebo time versus MADRS score curve was estimated.[51] This resulting MADRS difference was then converted to a utility and applied to all esketamine patients for one month and to those who partially or fully responded (i.e., continued esketamine beyond one month) for an

additional two months to reflect the QALYs gained by esketamine's rapid response in the first three-month cycle.

Table 4.7. Utility Values for Health States

Parameter	Base-Case Value	Source
No Depression **Male (not age specific)** **Female (not age specific)** **Gender Adjusted (not age specific)**	0.85 0.88 0.86	Sullivan 2006[94]
Mild to Moderate (weighted average of mild and moderate)	0.68	Janicak 2013[95]
Severe (weighted average of moderately severe and severe)	0.50	Janicak 2013[95]

Economic Inputs

Drug Utilization

The anticipated esketamine drug utilization is shown in Table 4.8. Initial dosing of esketamine in the TRANSFORM-2 study was either 56 mg or 84 mg twice weekly for one month.[51] In month two, patients received esketamine once weekly. From month three onward, patients were able to decrease the frequency of dosing to either once weekly or every other week. The proportions of patients taking each of the dosing strengths for initial and maintenance dosing frequencies are reported in Table 4.8.

Table 4.8. Treatment Regimen Recommended Dosage

Generic Name	Esketamine	Source
Brand Name	Spravato™	--
Manufacturer	Janssen	--
Route of Administration	Intranasal (clinic use only)	--
Proportion of Patients Receiving 56 or 84 mg During Initial Dosing	56 mg twice weekly (33% of patients) 84 mg twice weekly (67% of patients)	TRANSFORM-2[51]
Proportion of Patients Receiving 56 or 84 mg and Dosing Frequency During Maintenance Dosing	56 mg once weekly (14.2% of patients) 84 mg once weekly (28.7% of patients) 56 mg every other week (18.8% of patients) 84 mg every other week (38.2% of patients)	Weighted average of patients receiving 56 and 84 mg dose from TRANSFORM-2 and mean dosing frequency from Table 2 (averaging those who changed back and forth from weekly to every other week) in the SUSTAIN-2 trial.[51,54]
FDA-Approved Dosing Schedule	Induction (weeks 1-4): 56 or 84 mg twice weekly Maintenance (weeks 5-8): 56 or 84 mg once weekly Maintenance (weeks 5-8): 56 or 84 mg every other week	Spravato™ (esketamine) package insert[86]

Drug Costs

We used the wholesale acquisition cost (WAC) for pricing esketamine nasal spray in our analyses. Esketamine's unique mechanism of action among approved therapies for TRD, coupled with no current or anticipated competition in the therapeutic landscape of TRD which has a significant unmet treatment need led us to believe that any discounts or rebates for esketamine would likely be small. We thus applied its WAC price for our analyses. A WAC price of $295 per 28 mg device[97] was applied to the utilization doses and proportions of patients receiving each dose for esketamine (Table 4.8). Since esketamine requires observation of the patient for two hours after each administration, a physician office visit (CPT code 99214) was assigned for each dose, estimated using the Centers for Medicare and Medicaid Services (CMS) Physician Fee Schedule.[98]

Costs for alternative treatments used in the model were derived from data on pharmaceutical costs by number of depression medication regimen changes.[91] These costs from the year 2000 were then inflated to 2018 US dollars as per ICER's Reference Case. Since patients had failed a mean of three prior therapies in TRANSFORM-1 & -2, patients entered the model receiving a fourth depression medication regimen change plus esketamine or no additional therapy. The cost of esketamine was added to the underlying cost of the other depression medications. Each subsequent change in the antidepressant medication regiment (i.e., alternative treatments) resulted in increased pharmaceutical costs. These costs, labeled initial and alternative treatments in the model, are shown in Table 4.9.

Table 4.9. Alternative Treatment Costs

Current Number of Depression Medication Regimen Changes	Annual Cost (Inflated to 2018 USD)	Source
Initial Treatment, Not Including Esketamine (Fourth Regimen Change)	$3,909	Russell 2004[91]
First Alternative Treatment (Fifth Regimen Change)	$4,480	Russell 2004[91]
Second Alternative Treatment (Sixth Regimen Change)	$5,162	Russell 2004[91]
Third Alternative Treatment (Seventh Treatment Change)	$5,752	Russell 2004[91]

Health Care Utilization Costs

Non-drug depression related health care utilization and costs were derived from data on inpatient and outpatient costs by number of depression medication regimen changes, obtained from the same source as the pharmaceutical costs.[91] Costs in year 2000 were inflated as per ICER's Reference Case and are shown in Table 4.10.

Table 4.10. Inpatient and Outpatient Direct Medical Costs

Current Number of Depression Regimen Medication Changes	Annual Cost (Inflated to 2018 USD)	Source
Initial Treatment, Not Including Esketamine (Fourth Regimen Change)	$11,155	Russell 2004[91]
First Alternative Treatment (Fifth Regimen Change)	$12,888	Russell 2004[91]
Second Alternative Treatment (Sixth Regimen Change)	$13,717	Russell 2004[91]
Third Alternative Treatment (Seventh Treatment Change)	$14,344	Russell 2004[91]

Productivity Costs

Productivity was considered in a scenario analysis. Productivity was derived from a study evaluating patients with major depressive disorder who completed the Work Productivity and Activity Impairment Questionnaire: General Health (WPAI-GH) from the 2013 US National Health and Wellness Survey.[99] In this study, QWPAI-GH scores were stratified by PHQ-9 score. A cross-walk between PHQ-9 and MADRS scores was used to generate estimated work productivity losses for patients with treatment-resistant depression. With adequate treatment of treatment-resistant depression, we assumed that patients who experienced work productivity losses or impairment would regain the ability to be equally productive as those with a PHQ-9 score of 0-4. This work productivity gain was applied only to the proportion of patients who were employed at the time of

the study (i.e., 3,058 patients of the 6,997 patients who participated in the full study). Work productivity was inflated to December 2018 using the Organization for Economic Co-Operation and Development hourly earnings.[100]

Sensitivity Analyses

We conducted one-way sensitivity analyses on all model inputs to identify the impact of parameter uncertainty and key drivers of model outcomes. Probabilistic sensitivity analyses were also performed by jointly varying sensitive model parameters over 10,000 simulations, then calculating 95% credible range estimates for each model outcome based on the results. We also performed threshold analyses for drug costs across a range of incremental cost-effectiveness ratios (from $50,000 to $150,000 per QALY).

Scenario Analyses

A modified societal perspective was conducted including productivity gains for a portion of patients with effectively treated treatment-resistant depression.

Model Validation

We used several approaches to validate the model. First, we provided preliminary methods to manufacturers, patient groups, and clinical experts and results to patient groups, and clinical experts. Based on feedback from these groups on our methods, we refined them in the model. Second, we evaluated face validity of changes in results by varying model input parameters. We performed model verification for model calculations using internal reviewers. Finally, we will provide the manufacturer of esketamine an opportunity to review and comment on the most recent version of the model base case during the comment period for this report.

Model validation also included comparing our model and analyses to any similar previously published studies and analyses. We searched the literature to identify economic evaluations that were similar to our analysis, with comparable populations, settings, perspective, and treatments.

Most cost-effectiveness analysis of antidepressants or other therapies were conducted in patients with MDD. We found no prior economic evaluations estimating the cost-effectiveness of esketamine in patients with TRD. The only pertinent economic evaluations for TRD included treatment with ECT or TMS and are described below.

A US-specific cost-effectiveness analysis by Ross et al. evaluated seven treatment strategies for TRD, one without ECT and six with ECT with zero to five lines of therapy prior to ECT, in patients with TRD.[101] Like the ICER model, this model was also built from a health care sector perspective and measured QALYs and cost per QALY as key outcomes. However, unlike the ICER model, Ross et al.'s model had a substantially shorter four-year time horizon. For each treatment line, patients in Ross

et al.'s model could transition to remission, response or non-response, and from remission to relapse, and from relapse. Upon relapse or non-response, patients moved to a subsequent treatment line. Unlike in the ICER model, which had 3-month cycle lengths, patients in Ross et al.'s model could transition between health states every month. ECT efficacy was derived from meta-analyses of ECT observational data as well as meta-analyses of ECT trials. Antidepressant treatment efficacy estimates in Ross et al.'s model were derived from the STAR*D trial, with first line remission and response coming from a meta-analysis since these estimates from the STAR*D trial were substantially lower than what was observed in several meta-analyses. In the ICER model, treatment efficacy estimates for esketamine were informed by the TRANSFORM-1&-2 (using a meta-analyses) and SUSTAIN-1 trials. Alternative treatment efficacy estimates were derived from the STAR*D trial. Both models used similar utility estimates for no depression. Non-response, relapse and initiation health states in Ross et al.'s model had a utility value of 0.58 while the similar ICER model state, severe depression, 0.50. Remission in Ross et al.'s model had a utility of 0.72, compared with the ICER model's mild to moderate state utility of 0.68. Annual health care costs in both models are similar, from the fourth alternative treatment strategy onward, with these costs lower for the first three treatment alternatives in the ICER model compared to Ross et al.'s model. Although time horizons and treatment strategies were different in both models, Ross et al.'s model resulted in 2.63 (No ECT) and 2.76 (ECT as fourth-line treatment) QALYs over four years, for a difference of 0.12 QALYs gained. In contrast and because of the longer time horizon, the no additional treatment resulted in 12.64 QALYs, while treatment with esketamine resulted in 12.84 QALYs, for a difference of 0.20 QALYs gained. Cost per QALY gained findings were not compared due to the substantial differences in treatment options being compared.

A cost-effectiveness analysis by Zhao et al. compared repetitive TMS to ECT in patients with TRD in Singapore.[102] Treatment effectiveness estimates as well as health state utilities were derived from a local hospital database analysis as well as the published literature. The model included health care resources used in Singapore dollars, and reported outcomes as QALYs and incremental cost per QALY. Another study by Wiles et al. evaluated the cost-effectiveness of cognitive behavioral therapy as an adjunct to pharmacotherapy versus pharmacotherapy alone in patients with TRD in a UK-primary care setting.[103] This model used treatment efficacy and utility estimates from the CoBalT trial that was conducted across 73 primary care centers in the UK and was built from an NHS and personal social services perspective. Resource use was estimated from the trial and UK-specific costs applied. Since the above-mentioned models substantially differ from the ICER model in setting, treatments evaluated, and model estimates, a detailed comparison of modeling methods employed, and outcomes evaluated was not provided. These studies were included for reference only.

Cost-Analysis

A network meta-analysis comparing esketamine to ketamine was not possible due to substantial heterogeneity in the patients involved in these studies. A cost-analysis was conducted evaluating

the expected direct treatment costs for treatment with esketamine or ketamine. A *de novo* deterministic model was developed, informed by an analysis of resources used by intravenous ketamine clinics and anticipated resources used delivering intranasal esketamine in a clinic setting. Costs were applied to resources utilized, using published cost and fee structures. To estimate physician and clinic fees, we utilized the Calendar Year 2019 Medicare Physician Fee Schedule.[104] Supplies for intravenous drug administration were abstracted from the lowest available average wholesale prices from the McKesson Wholesale Medical Supply Ordering Platform (McKesson, San Francisco, CA). Labor costs for drug preparation were estimated using the Bureau of Labor Statistics.[105] We used the WAC for pricing esketamine and ketamine.[97] Prices of both drugs were applied to anticipated average annual usage for a patient continuing therapy with perfect adherence. Average annual usage was estimated using expert opinion for ketamine and clinical trials for esketamine[50,51] and is shown in table 4.8. A mean dose of intravenous ketamine 0.5 mg/kg, given six times in month one and then once monthly was used. For esketamine, an average dose of 74.8 mg given a mean of eight times in the first month, four times in the second month, and 2.86 times in the third and subsequent months.

4.3 Results

Base-Case Results

The main results are summarized in Tables 4.11 and 4.12. Given the base-case discontinuation rates, the model predicted that esketamine was being used by 19% of the initial cohort at three years, 4% at five years, and less than 1% by eight years. The results presented are hence reflective of treatment pathways that include initiation with esketamine or an oral antidepressant, and not just these initial treatments alone. The total discounted lifetime costs for esketamine and no additional treatment were $448,600 and $410,200, respectively. The total discounted lifetime QALYs in esketamine and no additional treatment arms were 12.66 and 12.47, respectively. The total discounted LYs gained were 20.66 (esketamine) and 20.64 (no additional treatment), respectively. This fractionally better survival in esketamine was due to the modeled impact of the treatment, which slows down progression to more severe depression states and subsequently results in a lower death rate from severe depression. The lifetime incremental cost-effectiveness ratio for esketamine compared with no additional treatment was approximately $198,000 per QALY gained. Cost per LY gained was $2.6 million and the cost per depression-free day was approximately $330 (over a two-year time horizon). All undiscounted cost and health outcomes are presented in Appendix Table E3.

Table 4.11. Base-Case Results Comparing Esketamine to No Additional Treatment in Patients with TRD

Treatment Pathways	Drug Cost	Total Cost	QALYs	LYs	Depression-Free Day
Esketamine	$42,600	$448,600	12.66	20.66	235
No Additional Treatment	$0	$410,200	12.47	20.64	117
Difference	$42,600	$38,400	0.19	0.01	117

QALY: quality-adjusted life year, LY: life year

Table 4.12. Incremental Cost-Effectiveness Ratios for the Base-Case Analysis

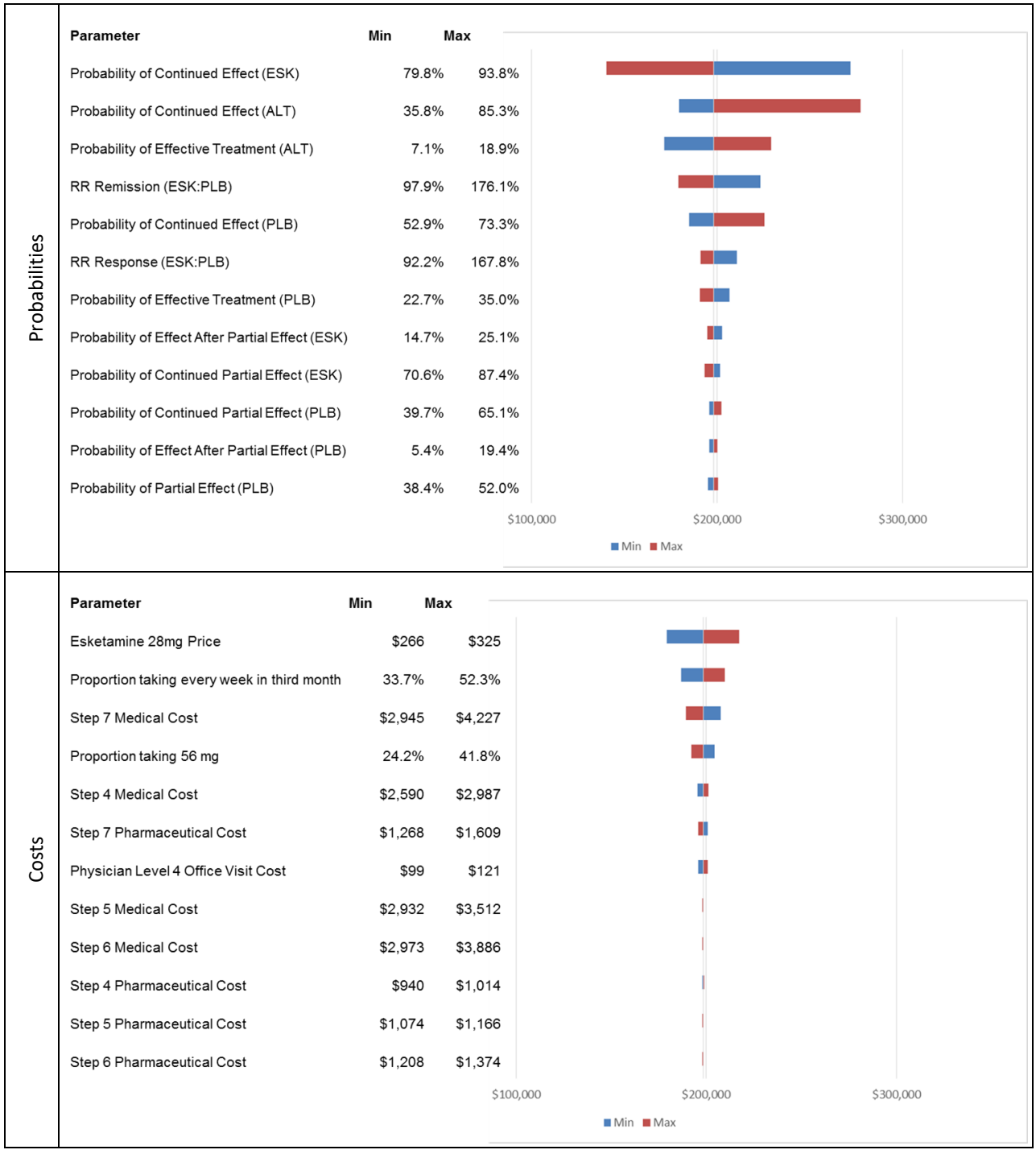
Treatment Pathways	Cost Per QALY Gained	Cost Per LY Gained	Cost Per Depression-Free Day
Esketamine vs. No Additional Treatment	$198,000	$2,592,000	$330

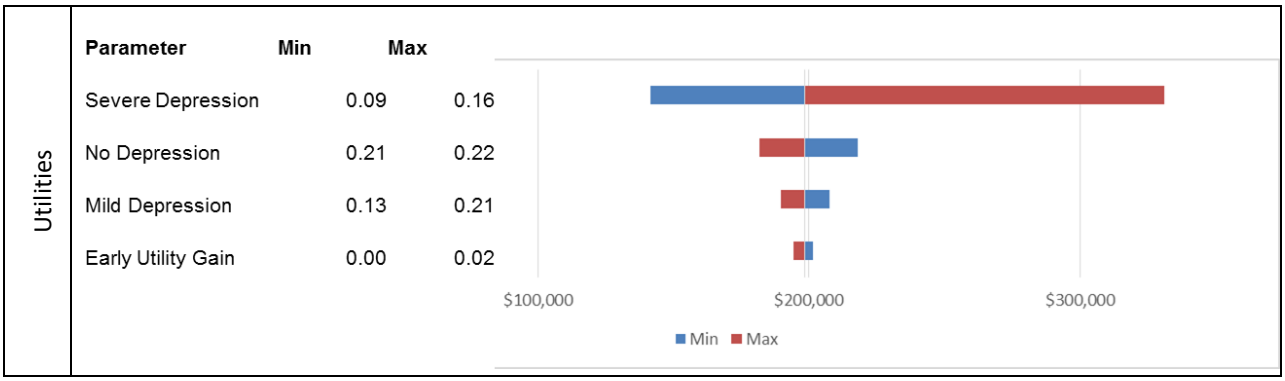
QALY: quality-adjusted life year, LY: life year

Sensitivity Analysis Results

To demonstrate effects of uncertainty on cost per QALY gained, we varied input parameters on reasonable ranges.

Figure 4.2. Tornado Diagrams for One-Way Sensitivity Analyses of Esketamine Versus No Additional Treatment in Patients with TRD





The results of the probabilistic sensitivity analysis are summarized in the table below and in Appendix Table E4. At willingness to pay thresholds of $150,000 per QALY gained or lower, treatment with esketamine was considered cost-effective in 15% or fewer of the 10,000 simulation runs. Treatment with esketamine became cost-effective in 50% of all simulation runs only at a WTP threshold of approximately $200,000 per QALY gained.

Table 4.13. Probabilistic Sensitivity Analysis Results

	Cost-Effective at $50,000 Per QALY	Cost-Effective at $100,000 Per QALY	Cost-Effective at $150,000 Per QALY
Esketamine	0%	1%	15%

Scenario Analyses Results

Modified Societal Perspective

When labor benefits for the proportion of patients who worked were included in the analysis, the lifetime incremental cost-effectiveness ratio for esketamine compared with no additional treatment was $188,000 per QALY gained. Detailed outcomes for the modified societal perspective are presented in Appendix Table E5.

Threshold Analyses Results

Average price per 28mg nasal spray device that would result in willingness-to-pay thresholds of $50,000 to $150,000 per QALY gained are shown in table 4.14 below.

Table 4.14. Threshold Analysis Results

	WAC Per Unit 28 mg Device	Price Per 28 mg Device to Achieve $50,000 Per QALY	Price per 28 mg Device to Achieve $100,000 Per QALY	Price Per 28 mg Device to Achieve $150,000 Per QALY
Esketamine	$295	$64	$142	$220

QALY: quality-adjusted life year

Cost Analysis Results

For a patient continuing therapy for a full year and including all administration costs, the first year of esketamine treatment resulted in an estimated annual direct cost of approximately $36,500 compared with approximately $3,600 for ketamine treatment. The annual direct costs for year two and future years was estimated to be approximately $30,800 and $2,500, respectively. When indirect costs associated with lost time from work and travel to and from the clinic were included, the first-year cost for esketamine and ketamine were approximately $39,400 and $5,300. The second and future year annual costs, including indirect costs, were approximately $33,300 and $3,700, respectively.

4.4 Summary and Comment

In our analysis of the cost-effectiveness of esketamine plus a new oral antidepressant compared with no additional treatment beyond a new oral antidepressant in patients with TRD, we found that esketamine produces substantial gains in quality of life while patients are taking the drug, although few patients in the model continued esketamine beyond five years. At the base-case price of $295 per 28 mg intranasal device, esketamine use results in an incremental cost-effectiveness ratio of approximately $198,000 per QALY compared to no additional treatment, well above the commonly-cited cost-effectiveness threshold of $150,000 per QALY. Esketamine use also resulted in cost per LY gained of approximately $2.6 million relative to no additional therapy, which is largely due to the marginally better survival in the esketamine arm. The inclusion of productivity gains from improved mood did not result in treatment with esketamine meeting the $150,000 per QALY gained threshold. In one-way sensitivity analyses, the model was sensitive to the probabilities determining the continued effectiveness of esketamine, its comparator, or alternative treatment. Also, the remission rate ratio of esketamine compared to placebo as calculated from our meta-analysis was an important factor determining esketamine's cost-effectiveness ratio. The model was also sensitive to the price of esketamine and the utility associated with having severe depression.

Importantly, the place for esketamine may depend on the comparative benefits between esketamine and other available treatments, such as ketamine. Unfortunately, such information is not available at this time. The one-year costs of esketamine are substantially higher than those of ketamine, even when considering increased administration costs associated with providing ketamine intravenously. Finally, the effectiveness (and cost-effectiveness) of esketamine alone, without a change to the current antidepressant regimen, is not known at this time.

Limitations

This analysis has several limitations and assumptions that must be considered when evaluating the results. The analysis was limited by the lack of comparative effectiveness data of esketamine to other commonly used treatments for TRD as this analysis only compared esketamine plus a new

oral antidepressant to a placebo plus a new oral antidepressant. For example, ketamine is a commonly used alternative treatment for TRD and its inclusion in this analysis may have been more useful for decision makers as a placebo would not typically be considered a treatment option in practice. Possible treatment comparators such as ketamine were considered for inclusion in an NMA. However, due to limitations in study design and populations enrolled, it was not possible to conduct an NMA nor evaluate the relative costs and benefits of treatment with esketamine to other alternative treatments.

Treatment-resistant depression is often defined using the number of treatment failures in the current depression episode. However, it is likely that the effectiveness of therapy, along with the total costs of care, depend on the number of treatments failed during a person's lifetime, pattern and frequency of depression episodes, and severity of the episodes. Detailed data of these important modifiers and their effects on patient outcomes, costs and quality of life have not been well studied. Assumptions were needed to use available estimates in the model, affecting the model structure and parameters. For example, a thorough review of the literature revealed that cost estimates were not available for medical care stratified by disease severity. Therefore, the model was designed to incorporate costs by number of treatment regimens (lifetime), for which limited data did exist. However, the number of lifetime treatment regimens was not available for the TRANSFORM trial. We therefore had to assume that costs for the number of treatments in the current depression episode mirrored costs for lifetime treatment regimens. While these model parameters were tested using extensive sensitivity analyses, the base-case results are particularly susceptible to bias in these estimates. As more evidence becomes available of the impact of important disease modifiers on clinical outcomes, cost of care, and patient quality of life, the model structure and inputs can be updated to incorporate our better understanding of TRD

For our scenario analysis, conducted from a modified societal perspective, we included cost benefits resulting from increased productivity with improved depression. These estimates were obtained from a study that estimated patient-reported absenteeism and presenteeism resulting from depression in a working population. We did not include the effects of depression on underemployment (or reemployment with treated depression) in the model, as we could not identify whether treatment of depression impacts reemployment nor could we find estimates for the possible effect of treatment on employment.

Conclusions

Compared with no additional treatment beyond a new oral antidepressant, TRD treatment with esketamine plus a new oral antidepressant resulted in important gains in patient QALYs over the lifetime. However, at its current price, esketamine is not cost-effective even at a WTP threshold of $150,000 per QALY gained. The results of this analysis should be considered in the context of a lack of evidence surrounding the treatment of TRD, including the complete lack of comparative evidence of esketamine to other potential therapies (i.e., ketamine) and very limited evidence of the impact of important disease modifiers on clinical outcomes, cost of care, and patient quality of life.

5. Potential Other Benefits and Contextual Considerations

Our reviews seek to provide information on potential other benefits offered by the intervention to the individual patient, caregivers, the delivery system, other patients, or the public that would not have been considered as part of the evidence on comparative clinical effectiveness. We also recognize that there may be broader contextual issues related to the severity of the condition, whether other treatments are available, and ethical, legal, or other societal priorities that influence the relative value of illnesses and interventions. These general elements are listed in the table below, and the subsequent text provides detail about the elements that are applicable to the comparison of esketamine. We sought input from stakeholders, including individual patients, patient advocacy organizations, clinicians, and manufacturers, to inform the contents of this section.

Each ICER review culminates in a public meeting of an independent voting Council of clinicians, patients, and health services researchers. As part of their deliberations, Council members will judge whether a treatment may substantially impact the considerations listed in Table 5.1. The presence of substantial other benefits or contextual considerations may shift a council member's vote on an intervention's long-term value for money to a different category than would be indicated by the clinical evidence and cost-effectiveness analyses alone. For example, a council member may initially consider a therapy with an incremental cost-effectiveness ratio of $150,000 per QALY to represent low long-term value for money. However, the Council member may vote for a higher value category if they consider the treatment to bring substantial other benefits or contextual considerations. Conversely, disadvantages associated with a treatment may lead a Council member to vote for a lower value category. A Council member may also determine that there are no other benefits or contextual considerations substantial enough to shift their vote. All factors that are considered in the voting process are outlined in ICER's value assessment framework. The content of these deliberations is described in the last chapter of ICER's Final Evidence Report, which is released after the public meeting.

This section, as well as the Council's deliberation, provides stakeholders with information to inform their decisions on a range of issues, including shared decision-making between patients and clinicians, coverage policy development, and pricing negotiations.

Table 5.1. Potential Other Benefits or Contextual Considerations (Not Specific to Any Disease or Therapy)

Potential Other Benefits
This intervention offers reduced complexity that will significantly improve patient outcomes.
This intervention will reduce important health disparities across racial, ethnic, gender, socio-economic, or regional categories.
This intervention will significantly reduce caregiver or broader family burden.
This intervention offers a novel mechanism of action or approach that will allow successful treatment of many patients for whom other available treatments have failed.
This intervention will have a significant impact on improving return to work and/or overall productivity.
Other important benefits or disadvantages that should have an important role in judgments of the value of this intervention.
Potential Other Contextual Considerations
This intervention is intended for the care of individuals with a condition of particularly high severity in terms of impact on length of life and/or quality of life.
This intervention is intended for the care of individuals with a condition that represents a particularly high lifetime burden of illness.
This intervention is the first to offer any improvement for patients with this condition.
Compared to background antidepressant alone there is significant uncertainty about the long-term risk of serious side effects of this intervention.
Compared to background antidepressant alone there is significant uncertainty about the magnitude or durability of the long-term benefits of this intervention.
There are additional contextual considerations that should have an important role in judgments of the value of this intervention.

5.1 Potential Other Benefits

For patient with TRD, esketamine is the first drug to receive FDA approval whose mechanism of action is thought to be through the NMDA receptor. Its development was based upon evidence that ketamine, a longstanding anesthetic drug, provides short-term improvement in mood and depressive symptoms. As a therapy that offers a novel mechanism of action, esketamine presents an alternative option for those patients with TRD who do not find relief or suffer severe side effects from other available treatments.

Due to the need for intranasal administration, esketamine may result in increased health care complexity. In addition, esketamine will be made available only through Risk Evaluation and Mitigation Strategy (REMS) program that will require dosing of the medicine in an approved doctor's office or clinic and monitoring by a health care provider for at least two hours after administration. This makes esketamine considerably more complex to administer and monitor than oral antidepressant medicines. However, for patients who have failed multiple oral medications, the burden of using esketamine needs to be considered in relationship to other commonly considered options. Ketamine is currently used as an off-label treatment primarily by IV infusion at

clinics, so the burden would be similar and even greater than for esketamine given the need to establish an IV. Other treatments may also include rTMS and ECT, both of which involve considerable logistical efforts. Even psychotherapy requires regular visits, and while it may involve less total time in the office, there is still the travel to and from the visit.

For patients who have had chronic, treatment-resistant MDD, the burden of this condition can result in a profound impact upon quality of life. This includes relationships with family and friends, ability to participate in educational and work activities, and even perform activities of daily living. The availability of a drug from a novel medicine class may provide patients with disabling MDD and their providers an important new option when existing medicines, psychotherapy and other treatments have failed or cannot be tolerated due to side effects.

It is unclear how esketamine will affect racial, ethnic, gender, socio-economic, or regional disparities. If the cost of treatment is significant, those with limited financial resources may find it difficult to afford treatment. Lack of access to high quality care for those with MDD may also play a role in poor diagnosis and management overall. Though patients and advocates expressed interest in new therapies for patients with TRD, they were cautious about how important an advance this would be given the nature of its dosing and administration. Thus, it is unclear if the introduction of esketamine will be viewed as addressing the need for new treatment options for those with this common, debilitating condition.

5.2 Contextual Considerations

Esketamine represents the first drug with a new mechanism of action for depression approved by the FDA in many years. The arrival of any new treatment option is seen as a positive development for those suffering from a chronic disease such as major depressive disorder. Patients and clinicians expressed interest in having new treatment options available for those with TRD.

Esketamine may be most appropriate for patients with TRD that is severe in nature and who have not responded to or tolerated multiple other therapies. On the other hand, for those with milder symptoms or having failed only a single therapy, the benefits and risks of esketamine may argue for other therapies first, as these patients were not included in the esketamine trials. Even patients with moderate symptoms who failed two other treatments in the current episode may consider other oral medications or psychotherapy prior to considering esketamine. This may reflect uncertainty about the comparative benefit of esketamine versus other treatments that may not have yet been tried, especially given the lack of long-term data.

For any new medication that has mainly been evaluated in short-term comparative trials, the long-term benefits and harms of esketamine are uncertain relative to other therapies that have years of experience. For patients who improve with esketamine and have tolerable side effects, it is uncertain how long to treat them for. Studies suggest that discontinuing esketamine is associated

with a higher rate of relapse than continuing it. The question then is how long it should be used and what are its long-term benefits and harms, especially compared to other treatment options.

Available data suggests that patients can remain on weekly or every other week esketamine for up to a year. However, longer term use and the potential for side effects not seen during short-term use remain. For example, use of esketamine is associated with transient side effects with dosing such as dissociation and elevated blood pressure. With longer term use, it is unclear if side effects not seen in short-term studies such as misuse or increased cardiovascular events may be observed. This may be a particular concern for patients with a history of substance use disorder or in elderly patients.

Specific subgroups of patients with TRD that are commonly encountered were excluded in the studies of esketamine. Available studies have not evaluated the use of esketamine in individuals with depression who also have acute suicidal ideation, psychosis, bipolar disorder (termed depression with mixed features), substance use disorders or anxiety disorders. In phase II and III trials, there were three deaths due to suicide in patients receiving esketamine and none in those receiving placebo. In addition, data has not been presented from the studies stratified by the severity of baseline symptoms (e.g., moderate or severe), the duration of the episode (e.g., greater or less than 1 year) or the number of years that the patient has had MDD. Whether esketamine is effective and safe in such subgroups of those with TRD is unknown.

Even for those who derive benefit from esketamine, the need for frequent dosing in a clinician's office with the need to monitor the patient for up to one and a half hours and then not drive a motor vehicle for the rest of day means that treatment administration and travel may take up at least half a day. Thus, while the benefits of esketamine may permit a patient to maintain employment or return to work, it may still result in missed time from work that is hard to explain or may slow advancement. Finally, patients and patient advocates expressed concern about the potential high cost of esketamine. Even when covered by health insurance, out of pocket costs can remain considerable and may prevent access to those who may benefit from esketamine.

6. Value-Based Price Benchmarks

Value-based price benchmarks will be included in the revised Evidence Report that will be released on or about May 9, 2019.

7. Potential Budget Impact

7.1 Overview

We used the cost-effectiveness model to estimate the potential total budgetary impact of treatment with esketamine plus a new oral antidepressant versus no additional treatment in adults diagnosed with treatment-resistant depression in the US. As in the cost-effectiveness model, treatment-resistance was defined as non-response to two or more adequate trials of antidepressant treatment in the current depressive episode.[5] Esketamine's unique mechanism of action among approved therapies for TRD, coupled with no current or anticipated competition in the therapeutic landscape of TRD which has a significant unmet treatment need led us to believe that any discounts or rebates for esketamine would likely be small. We therefore applied its WAC price in addition to the three threshold prices ($50,000, $100,000 and $150,000 per QALY) for esketamine in our estimates of budget impact.

7.2 Methods

We used results from the same model employed for the cost-effectiveness analyses to estimate total potential budget impact. Potential budget impact was defined as the total differential cost of using the new therapy in addition to relevant existing therapy for the treated population, calculated as differential health care costs (including drug costs) minus any offsets in these costs from averted health care events. All costs were undiscounted and estimated over a five-year time horizon. The five-year timeframe was of interest, given the potential for cost offsets to accrue over time and to allow a more realistic impact on the number of patients treated with the new therapy.

To estimate the size of the potential candidate population for treatment, we first identified the prevalence of MDD among adults in the US (7.1%). This estimate was based on results of the 2017 National Survey on Drug Use and Health (NSDUH) conducted by Substance Abuse and Mental Health Services Administration (SAMHSA).[106] While SAMHSA also reported data on those diagnosed and possibly treated for MDD (with or without medication), we could not derive the percentage of those formally diagnosed with MDD since this wasn't reported separately. We therefore applied an estimate of the percentage of those with MDD who were on treatment with an antidepressant/anti-psychotic medication (38.1%) to the prevalence estimate of MDD. This, along with the estimate on the prevalence of treatment-resistant depression (13.6%) among those with MDD, was derived from a US claims analysis.[107] Applying these filters to the 2019-2023 projected five-year average US adult population[108] resulted in an eligible population size of approximately 960,000 patients over five years, or approximately 192,000 patients each year who could be treated with esketamine plus an antidepressant in place of an antidepressant alone.

ICER's methods for estimating potential budget impact are described in detail elsewhere[109] and have been recently updated. The intent of our revised approach to budgetary impact is to document the percentage of patients who could be treated at specific prices without crossing a budget impact threshold that is aligned with overall growth in the US economy. For 2018-19, the five-year annualized potential budget impact threshold that should trigger policy actions to manage access and affordability is calculated to total approximately $991 million per year for new drugs.

To estimate potential budget impact, we evaluate a new therapy that would take market share from one or more existing therapies or treatments and calculate the blended budget impact associated with displacing use of existing therapies with the new intervention. For this analysis, since most patients on TRD are on an antidepressant or antipsychotic medication, we assumed all TRD patients were eligible for treatment with esketamine.

7.3 Results

Table 7.1 illustrates the average five-year annualized per-patient budget impact calculations for esketamine plus a new oral antidepressant compared to a new oral antidepressant alone in more detail, based on WAC ($295 per 28mg device) and the prices to reach $150,000, $100,000, and $50,000 per QALY for esketamine ($220, $142, and $64 per 28mg device, respectively).

Table 7.1. Annualized Per-Patient Budget Impact Calculations Over a Five-year Time Horizon

	Average Five-Year Annualized Per Patient Budget Impact			
	WAC	Price to Achieve $150,000 Per QALY	Price to Achieve $100,000 Per QALY	Price to Achieve $50,000 Per QALY
Esketamine + New Oral Antidepressant	$30,900	$27,900	$24,800	$21,700
New Oral Antidepressant Alone	$18,200			
Difference	$12,700	$9,700	$6,600	$3,500

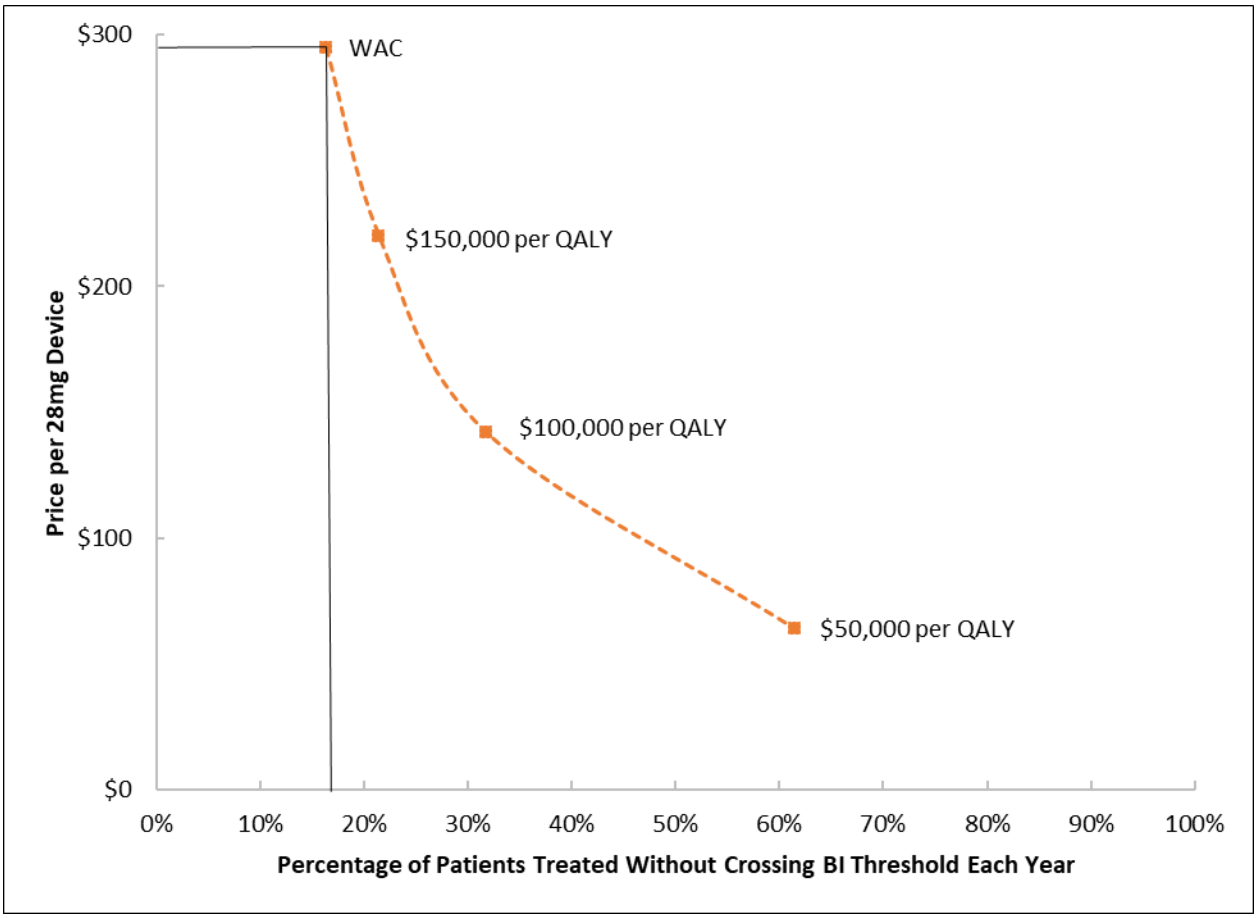
QALY: quality-adjusted life year, WAC: wholesale acquisition cost

The average five-year annualized potential budgetary impact of using esketamine plus a new oral antidepressant at esketamine's WAC was an additional per-patient cost of approximately $12,700. Average five-year annualized potential budgetary impact at the three cost-effectiveness threshold prices for esketamine ranged from approximately $9,700 per patient using esketamine's $150,000 per QALY cost-effectiveness threshold price to approximately $3,500 using its $50,000 per QALY threshold price.

As shown in Figure 7.1, over the five-year time horizon, 16% of eligible patients each year could be treated before the total budget exceeds the ICER budget impact threshold of $991 million at esketamine's WAC. This assumes equal uptake over the five years (20% each year), with treatment

duration ranging from one year (for the year-five cohort) to five years (for the year-one cohort). At prices to achieve WTP thresholds of $150,000 to $50,000 per QALY, between 21% and 62% of the eligible population could be treated before exceeding the $991 million threshold per year.

Figure 7.1. Potential Budget Impact Scenarios at Different Prices of Esketamine in TRD Patients



****

This is the first ICER review of esketamine.

References

1. National Institute of Mental Health. Major Depression-Data Sources. 2017; https://www.nimh.nih.gov/health/statistics/major-depression.shtml#part_155033. Accessed October 24, 2018.
2. Murray CJ, Atkinson C, Bhalla K, et al. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. *Jama.* 2013;310(6):591-608.
3. Ionescu DF, Rosenbaum JF, Alpert JE. Pharmacological approaches to the challenge of treatment-resistant depression. *Dialogues in clinical neuroscience.* 2015;17(2):111-126.
4. Souery D, Papakostas GI, Trivedi MH. Treatment-resistant depression. *The Journal of clinical psychiatry.* 2006;67 Suppl 6:16-22.
5. McIntyre RS, Filteau MJ, Martin L, et al. Treatment-resistant depression: definitions, review of the evidence, and algorithmic approach. *Journal of affective disorders.* 2014;156:1-7.
6. Nierenberg AA, DeCecco LM. Definitions of antidepressant treatment response, remission, nonresponse, partial response, and other relevant outcomes: a focus on treatment-resistant depression. *The Journal of clinical psychiatry.* 2001;62 Suppl 16:5-9.
7. Mrazek DA, Hornberger JC, Altar CA, Degtiar I. A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. *Psychiatric services (Washington, DC).* 2014;65(8):977-987.
8. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).* Arlington, VA: American Psychiatric Association; 2013.
9. Gartlehner G, Hansen RA, Morgan LC, et al. AHRQ Comparative Effectiveness Reviews. In: *Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2011.
10. Kennedy SH, Lam RW, McIntyre RS, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 3. Pharmacological Treatments. *Canadian journal of psychiatry Revue canadienne de psychiatrie.* 2016;61(9):540-560.
11. McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clinical psychology review.* 2009;29(3):243-259.
12. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American journal of psychiatry.* 2006;163(11):1905-1917.
13. Lex H, Ginsburg Y, Sitzmann AF, Grayhack C, Maixner DF, Mickey BJ. Quality of life across domains among individuals with treatment-resistant depression. *Journal of affective disorders.* 2019;243:401-407.
14. Santaguida PL, MacQueen G, Keshavarz H, Levine M, Beyene J, Raina P. AHRQ Comparative Effectiveness Reviews. In: *Treatment for Depression After Unsatisfactory Response to SSRIs.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.
15. Papadimitropoulou K, Vossen C, Karabis A, Donatti C, Kubitz N. Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis. *Current medical research and opinion.* 2017;33(4):701-711.

16. Milev RV, Giacobbe P, Kennedy SH, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 4. Neurostimulation Treatments. *Canadian journal of psychiatry Revue canadienne de psychiatrie.* 2016;61(9):561-575.
17. UK ECT Review Group. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. *Lancet (London, England).* 2003;361(9360):799-808.
18. Ren J, Li H, Palaniyappan L, et al. Repetitive transcranial magnetic stimulation versus electroconvulsive therapy for major depression: a systematic review and meta-analysis. *Progress in neuro-psychopharmacology & biological psychiatry.* 2014;51:181-189.
19. Trivedi RB, Nieuwsma JA, Williams JW, Jr. Examination of the utility of psychotherapy for patients with treatment resistant depression: a systematic review. *Journal of general internal medicine.* 2011;26(6):643-650.
20. Caddy C, Amit BH, McCloud TL, et al. Ketamine and other glutamate receptor modulators for depression in adults. *The Cochrane database of systematic reviews.* 2015(9):Cd011612.
21. Sanacora G, Frye MA, McDonald W, et al. A Consensus Statement on the Use of Ketamine in the Treatment of Mood Disorders. *JAMA psychiatry.* 2017;74(4):399-405.
22. Woolf S. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale. *Clinical Practice Guideline Development: Methodology Perspectives AHCPR Pub.* 1994:105-113.
23. Hamilton M. A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry.* 1960;23:56-62.
24. Hamilton M. Assessment of change in psychiatric state by means of rating scales. *Proceedings of the Royal Society of Medicine.* 1966;59 Suppl:10-13.
25. Hamilton M. Development of a rating scale for primary depressive illness. *The British journal of social and clinical psychology.* 1967;6(4):278-296.
26. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *The British journal of psychiatry : the journal of mental science.* 1979;134:382-389.
27. Åsberg M, Montgomery SA, Perris C, Schalling D, Sedvall G. A comprehensive psychopathological rating scale. *Acta psychiatrica Scandinavica Supplementum.* 1978(271):5-27.
28. Duru G, Fantino B. The clinical relevance of changes in the Montgomery-Åsberg Depression Rating Scale using the minimum clinically important difference approach. *Current medical research and opinion.* 2008;24(5):1329-1335.
29. Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Åsberg and the Hamilton ratings of depression: a comparison of measures. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology.* 2006;16(8):601-611.
30. Leucht S, Fennema H, Engel RR, Kaspers-Janssen M, Szegedi A. Translating the HAM-D into the MADRS and vice versa with equipercentile linking. *Journal of affective disorders.* 2018;226:326-331.
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine.* 2001;16(9):606-613.
32. Sheehan D. *Anxiety disease.* New York: Scribner; 1983.
33. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. *International clinical psychopharmacology.* 1996;11 Suppl 3:89-95.
34. Centers for Medicare & Medicaid Services. Local Coverage Determination: Transcranial Magnetic Stimulation. 2018. Accessed February 19, 2019.

35. Cigna HealthSpring. Clinical Guidelines - Repetitive Transcranial Magnetic Stimulation. 2013; https://www.cigna.com/static/docs/medicare-2019/clinical-practice-guidelines-rtms.pdf. Accessed February 19, 2019.
36. HealthNet. Clinical Policy: Transcranial Magnetic Stimulation. 2019; https://www.healthnet.com/static/general/unprotected/pdfs/national/policies/TMS.pdf. Accessed February 19, 2019.
37. WellCare Health Plans. Transcranial Magnetic Stimulation. 2015. Accessed February 19, 2019.
38. Aetna. Electroconvulsive Therapy. 2018; http://www.aetna.com/cpb/medical/data/400_499/0445.html. Accessed February 19, 2019.
39. Anthem Blue Cross and Blue Shield. Behavioral Health Medical Necessity Criteria. 2011; https://www11.anthem.com/provider/ct/f3/s9/t1/pw_e175456.pdf?refe. Accessed February 19, 2019.
40. Aetna. Transcranial Magnetic Stimulation and Cranial Electrical Stimulation. 2018; http://www.aetna.com/cpb/medical/data/400_499/0469.html.
41. Anthem. Transcranial Magnetic Stimulation Medical Policy. 2018; https://www11.anthem.com/ca/medicalpolicies/policies/mp_pw_a047769.htm. Accessed February 19, 2019.
42. Blue Cross Blue Shield of Kansas City. Transcranial Magnetic Stimulation and Cranial Electrical Stimulation (CES) as a Treatment of Depression and Other Psychiatric/Neurologic Disorders. 2018. Accessed February 19, 2019.
43. American Psychiatric Association. *Practice Guideline of the Treatment of Patients with Major Depressive Disorder.* 2010.
44. National Institute for Health and Care Excellence (NICE). Depression in adults: recognition and management. 2018. Accessed February 20, 2019.
45. World Health Organization. *Thirty-seventh Meeting of the Expert Committee on Drug Dependence.* 2015.
46. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med.* 1997;126(5):376-380.
47. Higgins JP. Cochrane Collaboration Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. 2008.
48. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery (London, England).* 2010;8(5):336-341.
49. Ollendorf DA, Pearson SD. An integrated evidence rating to frame comparative effectiveness assessments for decision makers. *Medical care.* 2010;48(6 Suppl):S145-152.
50. Fedgchin M, Trivedi M, Daly E, et al. Randomized, Double-Blind Study of Fixed-Dosed Intranasal Esketamine Plus Oral Antidepressant vs. Active Control in Treatment-Resistant Depression. *9th Biennial Conference of the International Society for Affective Disorders and the Houston Mood Disorders Conference.* 2018.
51. Popova V, Daly E, Trivedi M, et al. Randomized, Double-Blind study of Flexibly-dosed Intranasal Esketamine Plus Oral Antidepressant vs. Active Control in Treatment-Resistant Depression. *Annual Meeting of the American Society of Clinical Psychopharmacology.* 2018.
52. Ochs-Ross R, Daly EJ, Zhang Y, et al. Efficacy and safety of intranasal esketamine plus an oral antidepressant in elderly patients with treatment-resistant depression. *Annual Meeting of the American Society of Clinical Psychopharmacology.* 2018.
53. Daly E, Trivedi M, Janik A, et al. A Randomized Withdrawal, Double-blind, Multicenter Study of Esketamine Nasal Spray Plus an Oral Antidepressant for Relapse Prevention in Treatment-

resistant Depression. *Annual Meeting of the American Society of Clinical Psychopharmacology (ASCP).* 2018.
54. Wajs E, Aluisio L, Morrison R, et al. Long-Term Safety of Esketamine Nasal Spray Plus Oral Antidepressant in Patients with Treatment-Resistant Depression: Phase 3, Open-Label, Safety and Efficacy Study (SUSTAIN-2). *Annual Meeting of the American Society of Clinical Psychopharmacology.* 2018.
55. Singh JB, Fedgchin M, Daly EJ, et al. A double-blind, randomized, placebo-controlled, dose-frequency study of intravenous ketamine in patients with treatment-resistant depression. *The American journal of psychiatry.* 2016;173(8):816-826.
56. Murrough JW, Perez AM, Pillemer S, et al. Rapid and longer-term antidepressant effects of repeated ketamine infusions in treatment-resistant major depression. *Biological Psychiatry.* 2013;74(4):250-256.
57. Berman RM, Marcus RN, Swanink R, et al. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a multicenter, randomized, double-blind, placebo-controlled study. *The Journal of clinical psychiatry.* 2007;68(6):843-853.
58. Berman RM, Fava M, Thase ME, et al. Aripiprazole augmentation in major depressive disorder: a double-blind, placebo-controlled study in patients with inadequate response to antidepressants. *CNS spectrums.* 2009;14(4):197-206.
59. Marcus RN, McQuade RD, Carson WH, et al. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a second multicenter, randomized, double-blind, placebo-controlled study. *Journal of clinical psychopharmacology.* 2008;28(2):156-165.
60. Kamijima K, Higuchi T, Ishigooka J, et al. Aripiprazole augmentation to antidepressant therapy in Japanese patients with major depressive disorder: A randomized, double-blind, placebo-controlled study (ADMIRE study). *Journal of affective disorders.* 2013;151(3):899-905.
61. Kamijima K, Kimura M, Kuwahara K, Kitayama Y, Tadori Y. Randomized, double-blind comparison of aripiprazole/sertraline combination and placebo/sertraline combination in patients with major depressive disorder. *Psychiatry and Clinical Neurosciences.* 2018;72(8):591-601.
62. Thase ME, Youakim JM, Skuban A, et al. Efficacy and safety of adjunctive brexpiprazole 2 mg in major depressive disorder: a phase 3, randomized, placebo-controlled study in patients with inadequate response to antidepressants. *Journal of Clinical Psychiatry.* 2015;76(9):1224-1231.
63. Thase ME, Youakim JM, Skuban A, et al. Adjunctive brexpiprazole 1 and 3 mg for patients with major depressive disorder following inadequate response to antidepressants: a phase 3, randomized, double-blind study. *Journal of Clinical Psychiatry.* 2015;76(9):1232-1240.
64. Bauer M, Hefting N, Lindsten A, Josiassen MK, Hobart M. A randomised, placebo-controlled 24-week study evaluating adjunctive brexpiprazole in patients with major depressive disorder. *Acta Neuropsychiatrica.* 2018.
65. Hobart M, Skuban A, Zhang P, et al. A randomized, placebo-controlled study of the efficacy and safety of fixed-dose brexpiprazole 2 mg/d as adjunctive treatment of adults with major depressive disorder. *Journal of Clinical Psychiatry.* 2018;79(4).
66. Hobart M, Skuban A, Zhang P, et al. Efficacy and safety of flexibly dosed brexpiprazole for the adjunctive treatment of major depressive disorder: a randomized, active-referenced, placebo-controlled study. *Current medical research and opinion.* 2018;34(4):633-642.
67. Li X, Xing B, Yu E, Chen W, Wu H. The combined treatment of venlafaxine and quetiapine for treatment-resistant depression: a clinical study. *J Neuropsychiatry Clin Neurosci.* 2013;25(2):157-160.
68. Corya SA, Williamson D, Sanger TM, Briggs SD, Case M, Tollefson G. A randomized, double-blind comparison of olanzapine/fluoxetine combination, olanzapine, fluoxetine, and venlafaxine in treatment-resistant depression. *Depression and anxiety.* 2006;23(6):364-372.

69. Shelton RC, Williamson DJ, Corya SA, et al. Olanzapine/fluoxetine combination for treatment-resistant depression: a controlled study of SSRI and nortriptyline resistance. *The Journal of clinical psychiatry.* 2005;66(10):1289-1297.
70. George MS, Lisanby SH, Avery D, et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Archives of general psychiatry.* 2010;67(5):507-516.
71. Carpenter LL, Aaronson ST, Clarke GN, et al. rTMS with a two-coil array: Safety and efficacy for treatment resistant major depressive disorder. *Brain Stimulation.* 2017;10(5):926-933.
72. Brunelin J, Jalenques I, Trojak B, et al. The efficacy and safety of low frequency repetitive transcranial magnetic stimulation for treatment-resistant depression: The results from a large multicenter French RCT. *Brain Stimulation.* 2014;7(6):855-863.
73. Fitzgerald PB, Hoy KE, Herring SE, et al. A double blind randomized trial of unilateral left and bilateral prefrontal cortex transcranial magnetic stimulation in treatment resistant major depression. *Journal of affective disorders.* 2012;139(2):193-198.
74. Blumberger DM, Mulsant BH, Fitzgerald PB, et al. A randomized double-blind sham-controlled comparison of unilateral and bilateral repetitive transcranial magnetic stimulation for treatment-resistant major depression. *The world journal of biological psychiatry : the official journal of the World Federation of Societies of Biological Psychiatry.* 2012;13(6):423-435.
75. O'Reardon JP, Solvason HB, Janicak PG, et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry.* 2007;62(11):1208-1216.
76. Theleritis C, Sakkas P, Paparrigopoulos T, et al. Two versus one high-frequency repetitive transcranial magnetic stimulation session per day for treatment-resistant depression: A randomized sham-controlled trial. *The Journal of ECT.* 2017;33(3):190-197.
77. Bakim B, Uzun UE, Karamustafalioglu O, et al. The Combination of Antidepressant Drug Therapy and High-Frequency Repetitive Transcranial Magnetic Stimulation in Medication-Resistant Depression. *Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology.* 2012;22(3):244-253.
78. Chen SJ, Chang CH, Tsai HC, Chen ST, Lin CC. Superior antidepressant effect occurring 1 month after rTMS: Add-on rTMS for subjects with medication-resistant depression. *Neuropsychiatric Disease and Treatment.* 2013;9:397-401.
79. Pallanti S, Bernardi S, Di Rollo A, Antonini S, Quercioli L. Unilateral low frequency versus sequential bilateral repetitive transcranial magnetic stimulation: is simpler better for treatment of resistant depression? *Neuroscience.* 2010;167(2):323-328.
80. Blumberger DM, Maller JJ, Thomson L, et al. Unilateral and bilateral MRI-targeted repetitive transcranial magnetic stimulation for treatment-resistant depression: A randomized controlled study. *Journal of Psychiatry & Neuroscience.* 2016;41(4):E58-E66.
81. Zhang XH, Wang LW, Wang JJ, Liu Q, Fan Y. Adjunctive treatment with transcranial magnetic stimulation in treatment resistant depression: a randomized, double-blind, sham-controlled study. *Shanghai Archives of Psychiatry.* 2011;23(1):17-24.
82. Rosa MA, Gattaz WF, Pascual-Leone A, et al. Comparison of repetitive transcranial magnetic stimulation and electroconvulsive therapy in unipolar non-psychotic refractory depression: a randomized, single-blind study. *The international journal of neuropsychopharmacology.* 2006;9(6):667-676.
83. Food and Drug Administration. *Psychopharmacologic Drugs Advisory Committee (PDAC) and Drug Safety and Risk Management (DSaRM) Advisory Committee Meeting: Esketamine.* 2019.
84. Janssen Research & Development. *Advisory Committee Briefing Document: Esketamine Nasal Spray for Patients with Treatment-resistant Depression.* 2019.

85. Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical care.* 2004;42(12):1194-1201.
86. Food and Drug Administration (FDA). Spravato (Esketamine) Label. 2019; http://www.janssenlabels.com/package-insert/product-monograph/prescribing-information/SPRAVATO-pi.pdf.
87. Sassano-Higgins S, Baron D, Juarez G, Esmaili N, Gold M. A REVIEW OF KETAMINE ABUSE AND DIVERSION. *Depression and anxiety.* 2016;33(8):718-727.
88. Food and Drug Administration (FDA). Ketalar (Ketamine) Label. 2012; https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/016812s039lbl.pdf.
89. (FDA) FaDA. Electroconvulsive Therapy (ECT) Devices for Class II Intended Uses. 2015; https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM478942.pdf. Accessed March 19, 2019.
90. Daly EJ, Singh JB, Fedgchin M, et al. Efficacy and safety of intranasal esketamine adjunctive to oral antidepressant therapy in treatment-resistant depression: A randomized clinical trial. *JAMA psychiatry.* 2018;75(2):139-148.
91. Russell JM, Hawkins K, Ozminkowski RJ, et al. The cost consequences of treatment-resistant depression. *The Journal of clinical psychiatry.* 2004;65(3):341-347.
92. Human Mortality Database. 2016. www.mortality.org. Accessed February 2, 2019.
93. Reutfors J, Andersson TM, Brenner P, et al. Mortality in treatment-resistant unipolar depression: A register-based cohort study in Sweden. *Journal of affective disorders.* 2018;238:674-679.
94. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. *Medical decision making : an international journal of the Society for Medical Decision Making.* 2006;26(4):410-420.
95. Janicak PG, Dunner DL, Aaronson ST, et al. Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of quality of life outcome measures in clinical practice. *CNS spectrums.* 2013;18(6):322-332.
96. Chow W, Doane MJ, Sheehan J. Economic burden among patients with major depressive disorder: an analysis of health care resource use, work productivity, and direct and indirect costs by depression severity. Poster presented at: Academy of Managed Care Pharmacy NEXUS meeting; October 22-25, 2018; Orlando, FL. Poster F15. In.
97. Redbook. 2019. Accessed March 12, 2019.
98. Center for Medicare and Medicaid Services (CMS). Cost (national average price for non-facility and facility). source: CMS Physician fee Schedule. 2019; https://www.cms.gov/apps/physician-fee-schedule/license-agreement.aspx. Accessed March 19, 2019.
99. Chow W, Doane MJ, Sheehan J, Alphs L, Le H. Economic burden among patients with major depressive disorder: An analysis of health care resource use, work productivity, and direct and indirect costs by depression severity. *Am J Manag Care.* 2019;Suppl.(Feb):1-3 (Available at: https://www.ajmc.com/journals/supplement/2019/economic-burden-mdd-analysis-healthcare/economic-burden-mdd).
100. Hourly Earnings (MEI). 2019. https://stats.oecd.org/Index.aspx?DataSetCode=EAR_MEI. Accessed March 5, 2019.
101. Ross EL, Zivin K, Maixner DF. Cost-effectiveness of Electroconvulsive Therapy vs Pharmacotherapy/Psychotherapy for Treatment-Resistant Depression in the United States. *JAMA psychiatry.* 2018;75(7):713-722.
102. Zhao YJ, Tor PC, Khoo AL, Teng M, Lim BP, Mok YM. Cost-Effectiveness Modeling of Repetitive Transcranial Magnetic Stimulation Compared to Electroconvulsive Therapy for Treatment-Resistant Depression in Singapore. *Neuromodulation : journal of the International Neuromodulation Society.* 2018;21(4):376-382.

103. Wiles NJ, Thomas L, Turner N, et al. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial. *The lancet Psychiatry.* 2016;3(2):137-144.
104. Physician Fee Schedule Search. 2019. https://www.cms.gov/apps/physician-fee-schedule/search/search-criteria.aspx. Accessed January 27, 2019.
105. May 2017 National Occupational Employment and Wage Estimates United States. 2019. https://www.bls.gov/oes/current/oes_nat.htm#00-0000. Accessed March 19.
106. Center for Behavioral Health Statistics and Quality. *2017 National Survey on Drug Use and Health: Detailed Tables.* Rockville, MD: Substance Abuse and Mental Health Services Administration;2018.
107. Kubitz N, Mehra M, Potluri RC, Garg N, Cossrow N. Characterization of treatment resistant depression episodes in a cohort of patients from a US commercial claims database. *PloS one.* 2013;8(10):e76882.
108. United States Census Bureau. *2017 National Population Projections Tables.* Washington, D.C.2018.
109. Pearson SD. The ICER Value Framework: Integrating Cost Effectiveness and Affordability in the Assessment of Health Care Value. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2018;21(3):258-265.
110. Agency for Healthcare Research and Quality. U.S. Preventive Services Task Force Procedure Manual. 2008.
111. Sackeim HA. The definition and meaning of treatment-resistant depression. *The Journal of clinical psychiatry.* 2001;62 Suppl 16:10-17.
112. Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. *The Journal of clinical psychiatry.* 1997;58 Suppl 13:23-29.
113. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *Jama.* 2016;316(10):1093-1103.

APPENDICES

Appendix A. Search Strategies and Results

Table A1. PRISMA 2009 Checklist

	#	Checklist item
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
METHODS		
Protocol and Registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility Criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.
Information Sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study Selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data Collection Process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data Items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of Bias in Individual Studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary Measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of Results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.

	#	Checklist item
Risk of Bias Across Studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).
Additional Analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
RESULTS		
Study Selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study Characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of Bias Within Studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of Individual Studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of Results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of Bias Across Studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional Analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of Evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

From: Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

Table A2. Search Strategy of Medline 1996 to Present with Daily Update and PsychINFO Via Ovid

#	Search Terms
1	exp depressive disorder, treatment-resistant/
2	exp treatment resistant depression/
3	depress*.ti,ab.
4	(treatment-resist* or treatment resist* or therapy-resist* or therapy resist* or refract* or resist* or intractable or nonrespon* or non-respon* or unrespon* or fail* or ((no* or inadequat* or incomplet* or partial* or poor* or sub*) adj2 respon*) or (no* adj2 remi*)).ti,ab.
5	treatment failure/
6	drug resistance/
7	drug resistance, multiple/
8	4 or 5 or 6 or 7
9	3 and 8
10	1 or 2 or 9
11	(esketamine or S-ketamine or S ketamine or Ketanest or Ketanest S).ti,ab.
12	ketamine/
13	(ketamine or Ketaset or Ketalar or CI-581 or CI581 or CI 581 or Calipsol or Kalipsol or Calypsol or ketamin*).ti,ab.
14	(Citalopram or Escitalopram or Fluoxetine or Fluvoxamine or Paroxetine or Sertraline or Venlafaxine or Desvenlafaxine or Duloxetine or Milnacipran or Levomilnacipran or Amitriptyline or Amitriptylinoxide or Butriptyline or Clomipramine or Demexiptiline or Desipramine or Dibenzepin or Dimetacrine or Dosulepin or Dothiepin or Imipramine or Imipraminoxide or Lofepramine or Melitracen or Metapramine or Nitroxazepine or Nortriptyline or Noxiptiline or Pipofezine or Propizepine or Protriptyline or Quinupramine or Amineptine or Iprindole or Opipramol or Tianeptine or Trimipramine or Bupropion or Trazodone or Amoxapine or Maprotiline or Mazindol or Mianserin or Setiptiline or Isocarboxazid or Moclobemide or Phenelzine or Pirlindole or Selegiline or Tranylcypromine or Risperidone or Amisulpride or Aripiprazole or Asenapine or Clozapine or Iloperidone or Lurasidone or Olanzapine or Paliperidone or Quetiapine or Ziprasidone or Zotepine or Symbyax or Mirtazapin$ or Vortioxetine or Agomelatine or Doxepin or Reboxetine or Brexpiprazole or Vilazodone or Nefazodone).ti,ab.
15	Citalopram/ or Fluoxetine/ or Fluvoxamine/ or Paroxetine/ or Sertraline/ or Venlafaxine Hydrochloride/ or Desvenlafaxine Succinate/ or Duloxetine Hydrochloride/ or Milnacipran/ or Levomilnacipran/ or Amitriptyline/ or Clomipramine/ or Desipramine/ or Dothiepin/ or Imipramine/ or Lofepramine/ or Nortriptyline/ or Protriptyline/ or Iprindole/ or Opipramol/ or Trimipramine/ or Bupropion/ or Trazodone/ or Amoxapine/ or Maprotiline/ or Mazindol/ or Mianserin/ or Isocarboxazid/ or Moclobemide/ or Phenelzine/ or Selegiline/ or Tranylcypromine/ or Risperidone/ or Amisulpride/ or Aripiprazole/ or Clozapine/ or Lurasidone Hydrochloride / or Olanzapine/ or Paliperidone Palmitate/ or Quetiapine Fumarate/ or olanzapine-fluoxetine combination/ or Mirtazapine/ or Vortioxetine/ or Doxepin/ or Reboxetine/ or Vilazodone Hydrochloride/
16	(Celexa or Cipramil or Cipram or Dalsan or Recital or Emocal or Sepram or Seropram or Citox or Cital or Lexapro or Cipralex or Seroplex or Esertia or Depex or Prozac or Fontex or Seromex or Seronil or Sarafem or Ladose or Motivest or Flutop or Fluctin or Fluox or Lovan or Prodep or Luvox or Fevarin or Faverin or Dumyrox or Favoxil or Movox or Floxyfral or Paxil or Seroxat or Sereupin or Aropax or Deroxat or Divarius or Rexetin or Xetanor or Paroxat or Loxamine or Deparoc or Zoloft or Lustral or Serlain or Asentra or Tresleen or Effexor or Efexor or Cymbalta or Ariclaim or Xeristar or Yentreve or

#	Search Terms
	Duzela or Dulane or Wellbutrin or Budeprion or Prexaton or Elontril or Aplenzin or Risperdal or Parnate or Jatrosom or Tofranil or Tofranil-PM or Elavil or Endep or Vanatrip or Anafranil or Pamelor or Aventyl Hydrochloride or Desyrel or Oleptro or Beneficat or Deprax or Desirel or Molipaxin or Thombran or Trazorel or Trialodine or Trittico or Mesyrel or Meresa or Bosnyl or Dogmatil or Dolmatil or Eglonyl or Modal or Espiride or Abilify or Saphris or Sycrest or Leponex or Fanapt or Fanapta or Zomaril or Latuda or Zyprexa or Zalasta or Invega or Seroquel or Geodon or Zeldox or Pristiq or Dalcipran or Ixel or Savella or Fetzima or Tryptomer or Elavil or Endep or Amioxid or Ambivalon or Equilibrin or Evadyne or Deparon or Tinora or Norpramin or Pertofane or Noveril or Victoril or Istonil or Istonyl or Miroistonil or Prothiaden or Adapin or Sinequan or Tofranil or Janimine or Praminil or Imiprex or Elepsin or Lomont or Gamanil or Deanxit or Dixeran or Melixeran or Trausabun or Timaxel or Pamelor or Aventyl or Norpress or Agedal or Elronon or Nogedal or Azafen or Azaphen or Vagran or Vivactil or Kevopril or Kinupril or Adeprim or Quinuprine or Survector or Maneon or Directim or Prondol or Galatur or Tetran or Insidon or Pramolan or Ensidon or Oprimol or Stablon or Coaxil or Tatinol or Surmontil or Asendin or Deprilept or Ludiomil or Psymion or Mazanor or Sanorex or Tecipul or Marplan or Aurorix or Manerix or Nardil or Eldepryl or Zelapar or Emsam or Solian or Clozaril or Nipolept or Remergil or Remeron or Zispin or Remergon or Rexer or Promyrtil or Norset or Remeron SolTab or 6-Azamianserin or Mepirzepine or ORG-3770 or Brintellix or Valdoxan or Melitor or Thymanax or Deptran or Sinequan or Edronax or Prolift or OPC-34712 or Viibryd or Serzone).ti,ab.
17	electroconvulsive therapy/
18	(ect or electroconvulsive therapy).ti,ab.
19	transcranial magnetic stimulation/
20	(TMS or rTMS or transcranial magnetic stimulation).ti,ab.
21	14 or 15 or 16 or 17 or 18 or 19 or 20
22	limit 21 to yr="2013- Current"
23	11 or 12 or 13 or 22
24	10 and 23
25	(animals not (human and animals)).sh.
26	24 not 25
27	limit 26 to english language
28	(addresses OR autobiography OR bibliography OR biography OR case reports OR clinical trial, phase I OR comment OR congresses OR consensus development conference OR duplicate publication OR editorial OR guideline OR interview OR lectures OR legal cases OR legislation OR letter OR news OR newspaper article OR patient education handout OR periodical index OR personal narratives OR portraits OR practice guideline OR review OR video-audio media).pt
29	exp cohort studies/ OR comparative study.pt.
30	control groups/ or (control* adj2 (clinical or group* or trial* or study or studies or design* or arm*)).ti,ab. or (clinical trial or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or controlled clinical trial or multicenter study or randomized controlled trial).pt. or (randomi?ed adj6 (study or trial* or (clinical adj2 trial*))).ti,ab.
31	29 or 30
32	31 not 28
33	27 and 32
34	remove duplicates from 33
Date of search: December 5, 2018	

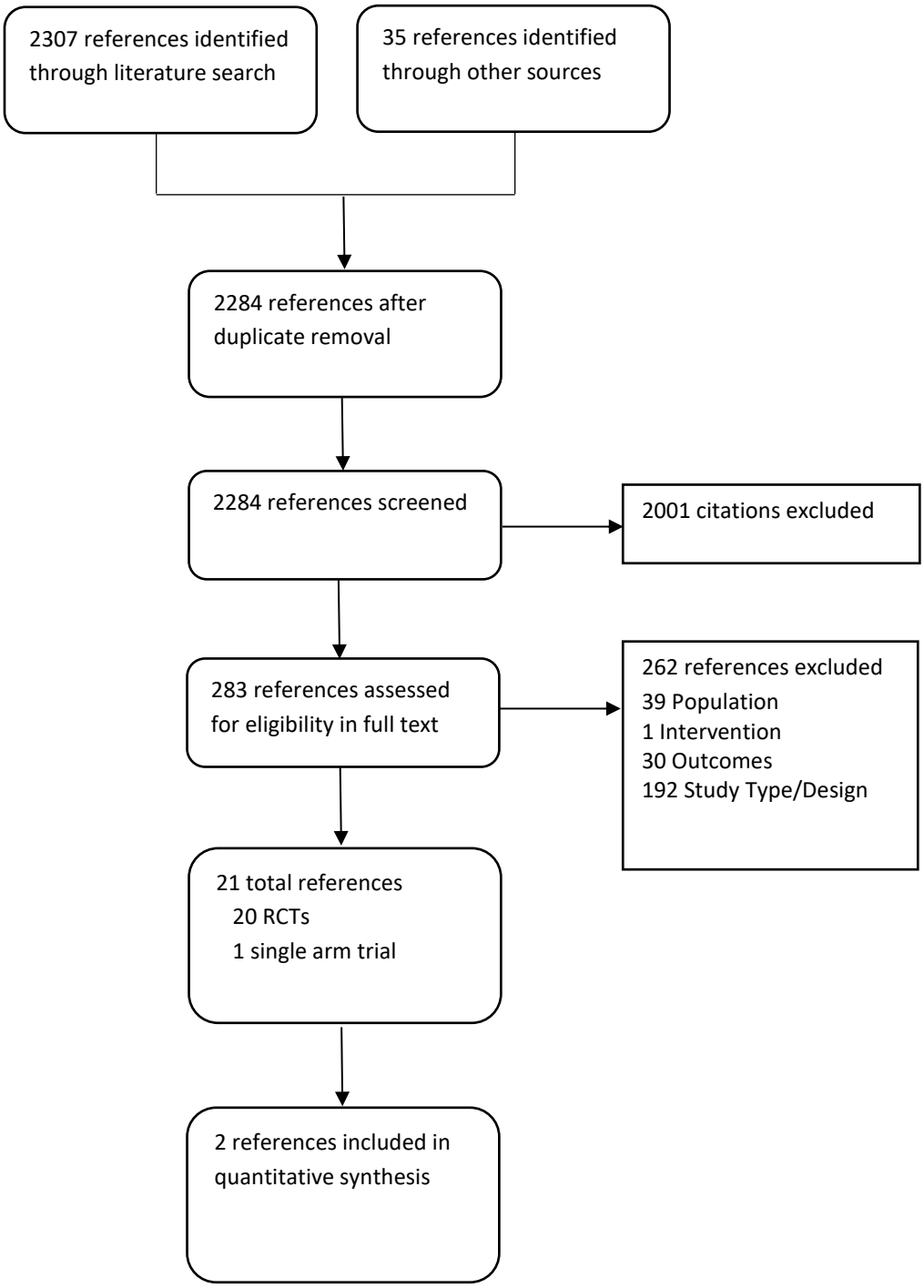
Table A3. Search Strategy of EMBASE

#	Search Terms
1	'treatment resistant depression'/exp
2	depress*:ti,ab
3	'treatment resist*':ti,ab OR 'treatment-resist*':ti,ab OR 'therapy resist*':ti,ab OR 'therapy-resist*':ti,ab OR refract*:ti,ab OR resist*:ti,ab OR intractable:ti,ab OR nonrespon*:ti,ab OR 'non-respon*':ti,ab OR fail*:ti,ab OR unrespon*:ti,ab OR (((no* OR inadequat* OR incomplet* OR partial* OR poor* OR sub*) NEAR/2 respon*):ti,ab) OR ((no* NEAR/2 remi*):ti,ab)
4	'treatment failure'/exp
5	'drug resistance'/de
6	'multidrug resistance'/de
7	#3 OR #4 OR #5 OR #6
8	#2 AND #7
9	#1 OR #8
10	'esketamine'/de
11	esketamine:ti,ab OR 's ketamine':ti,ab OR 's-ketamine':ti,ab OR ketanest:ti,ab
12	'ketamine'/de
13	ketamine:ti,ab OR ketaset:ti,ab OR ketalar:ti,ab OR ci581:ti,ab OR 'ci 581':ti,ab OR 'ci-581':ti,ab OR calipsol:ti,ab OR kalipsol:ti,ab OR calypsol:ti,ab or ketamin*:ti,ab
14	#10 OR #11 OR #12 OR #13
15	#9 AND #14
16	'clinical':ti,ab AND 'trial':ti,ab OR 'clinical trial'/exp OR random* OR 'drug therapy':lnk
17	'clinical article'/exp OR 'controlled study'/exp OR 'major clinical study'/exp OR 'prospective study'/exp OR 'cohort analysis'/exp OR 'cohort':ti,ab OR 'compared':ti,ab OR 'groups':ti,ab OR 'case control':ti,ab OR 'multivariate':ti,ab
18	#16 OR #17
19	#15 AND #18
20	'electroconvulsive therapy'/de
21	ect:ti,ab OR 'electroconvulsive therapy':ti,ab
22	'transcranial magnetic stimulation'/exp
23	tms:ti,ab OR rtms:ti,ab OR 'transcranial magnetic stimulation':ti,ab
24	citalopram:ti,ab OR escitalopram:ti,ab OR fluoxetine:ti,ab OR fluvoxamine:ti,ab OR paroxetine:ti,ab OR sertraline:ti,ab OR venlafaxine:ti,ab OR desvenlafaxine:ti,ab OR duloxetine:ti,ab OR milnacipran:ti,ab OR levomilnacipran:ti,ab OR amitriptyline:ti,ab OR amitriptylinoxide:ti,ab OR butriptyline:ti,ab OR clomipramine:ti,ab OR demexiptiline:ti,ab OR desipramine:ti,ab OR dibenzepin:ti,ab OR dimetacrine:ti,ab OR dosulepin:ti,ab OR dothiepin:ti,ab OR imipramine:ti,ab OR imipraminoxide:ti,ab OR lofepramine:ti,ab OR melitracen:ti,ab OR metapramine:ti,ab OR nitroxazepine:ti,ab OR nortriptyline:ti,ab OR noxiptiline:ti,ab OR pipofezine:ti,ab OR propizepine:ti,ab OR protriptyline:ti,ab OR quinupramine:ti,ab OR amineptine:ti,ab OR iprindole:ti,ab OR opipramol:ti,ab OR tianeptine:ti,ab OR trimipramine:ti,ab OR amfebutamone:ti,ab OR bupropion:ti,ab OR trazodone:ti,ab OR amoxapine:ti,ab OR maprotiline:ti,ab OR mazindol:ti,ab OR mianserin:ti,ab OR setiptiline:ti,ab OR teciptiline:ti,ab OR isocarboxazid:ti,ab OR moclobemide:ti,ab OR phenelzine:ti,ab OR pirlindole:ti,ab OR selegiline:ti,ab OR tranylcypromine:ti,ab OR risperidone:ti,ab OR amisulpride:ti,ab OR aripiprazole:ti,ab

#	Search Terms
	OR asenapine:ti,ab OR clozapine:ti,ab OR iloperidone:ti,ab OR lurasidone:ti,ab OR olanzapine:ti,ab OR paliperidone:ti,ab OR quetiapine:ti,ab OR ziprasidone:ti,ab OR zotepine:ti,ab OR 'fluoxetine near/1 olanzapine':ti,ab OR symbyax:ti,ab OR mirtazapin$:ti,ab OR vortioxetine:ti,ab OR agomelatine:ti,ab OR doxepin:ti,ab OR reboxetine:ti,ab OR brexpiprazole:ti,ab OR vilazodone:ti,ab OR nefazodone:ti,ab
25	'citalopram' OR 'escitalopram' OR 'fluoxetine' OR 'fluvoxamine' OR 'paroxetine' OR 'sertraline' OR 'venlafaxine' OR 'desvenlafaxine' OR 'duloxetine' OR 'milnacipran' OR 'levomilnacipran' OR 'amitriptyline' OR 'amitriptylinoxide' OR 'butriptyline' OR 'clomipramine' OR 'demexiptiline' OR 'desipramine' OR 'dibenzepin' OR 'dimetacrine' OR 'dosulepin' OR 'dothiepin' OR 'imipramine' OR 'imipraminoxide' OR 'lofepramine' OR 'melitracen' OR 'metapramine' OR 'nitroxazepine' OR 'nortriptyline' OR 'noxiptiline' OR 'pipofezine' OR 'propizepine' OR 'protriptyline' OR 'quinupramine' OR 'amineptine' OR 'iprindole' OR 'opipramol' OR 'tianeptine' OR 'trimipramine' OR 'amfebutamone' OR 'bupropion' OR 'trazodone' OR 'amoxapine' OR 'maprotiline' OR 'mazindol' OR 'mianserin' OR 'setiptiline' OR 'teciptiline' OR 'isocarboxazid' OR 'moclobemide' OR 'phenelzine' OR 'pirlindole' OR 'selegiline' OR 'tranylcypromine' OR 'risperidone' OR 'amisulpride' OR 'aripiprazole' OR 'asenapine' OR 'clozapine' OR 'iloperidone' OR 'lurasidone' OR 'olanzapine' OR 'paliperidone' OR 'quetiapine' OR 'ziprasidone' OR 'zotepin' OR 'zotepine' OR 'fluoxetine plus olanzapine' OR 'symbyax' OR 'mirtazapine' OR 'vortioxetine' OR 'agomelatine' OR 'doxepin' OR 'reboxetine' OR 'brexpiprazole' OR 'vilazodone' OR 'nefazodone'
26	celexa:ti,ab OR cipramil:ti,ab OR cipram:ti,ab OR dalsan:ti,ab OR recital:ti,ab OR emocal:ti,ab OR sepram:ti,ab OR seropram:ti,ab OR citox:ti,ab OR cital:ti,ab OR lexapro:ti,ab OR cipralex:ti,ab OR seroplex:ti,ab OR esertia:ti,ab OR depex:ti,ab OR prozac:ti,ab OR fontex:ti,ab OR seromex:ti,ab OR seronil:ti,ab OR sarafem:ti,ab OR ladose:ti,ab OR motivest:ti,ab OR flutop:ti,ab OR fluctin:ti,ab OR fluox:ti,ab OR lovan:ti,ab OR prodep:ti,ab OR luvox:ti,ab OR fevarin:ti,ab OR faverin:ti,ab OR dumyrox:ti,ab OR favoxil:ti,ab OR movox:ti,ab OR floxyfral:ti,ab OR paxil:ti,ab OR seroxat:ti,ab OR sereupin:ti,ab OR aropax:ti,ab OR deroxat:ti,ab OR divarius:ti,ab OR rexetin:ti,ab OR xetanor:ti,ab OR paroxat:ti,ab OR loxamine:ti,ab OR deparoc:ti,ab OR zoloft:ti,ab OR lustral:ti,ab OR serlain:ti,ab OR asentra:ti,ab OR tresleen:ti,ab OR effexor:ti,ab OR efexor:ti,ab OR cymbalta:ti,ab OR ariclaim:ti,ab OR xeristar:ti,ab OR yentreve:ti,ab OR duzela:ti,ab OR dulane:ti,ab OR wellbutrin:ti,ab OR budeprion:ti,ab OR prexaton:ti,ab OR elontril:ti,ab OR aplenzin:ti,ab OR risperdal:ti,ab OR parnate:ti,ab OR jatrosom:ti,ab OR 'tofranil pm':ti,ab OR vanatrip:ti,ab OR anafranil:ti,ab OR 'aventyl hydrochloride':ti,ab OR desyrel:ti,ab OR oleptro:ti,ab OR beneficat:ti,ab OR deprax:ti,ab OR desirel:ti,ab OR molipaxin:ti,ab OR thombran:ti,ab OR trazorel:ti,ab OR trialodine:ti,ab OR trittico:ti,ab OR mesyrel:ti,ab OR meresa:ti,ab OR bosnyl:ti,ab OR dogmatil:ti,ab OR dolmatil:ti,ab OR eglonyl:ti,ab OR modal:ti,ab OR espiride:ti,ab OR abilify:ti,ab OR saphris:ti,ab OR sycrest:ti,ab OR leponex:ti,ab OR fanapt:ti,ab OR fanapta:ti,ab OR zomaril:ti,ab OR latuda:ti,ab OR zyprexa:ti,ab OR zalasta:ti,ab OR invega:ti,ab OR seroquel:ti,ab OR geodon:ti,ab OR zeldox:ti,ab OR pristiq:ti,ab OR dalcipran:ti,ab OR ixel:ti,ab OR savella:ti,ab OR fetzima:ti,ab OR tryptomer:ti,ab OR elavil:ti,ab OR endep:ti,ab OR amioxid:ti,ab OR ambivalon:ti,ab OR equilibrin:ti,ab OR evadyne:ti,ab OR deparon:ti,ab OR tinora:ti,ab OR norpramin:ti,ab OR pertofane:ti,ab OR noveril:ti,ab OR victoril:ti,ab OR istonil:ti,ab OR istonyl:ti,ab OR miroistonil:ti,ab OR prothiaden:ti,ab OR adapin:ti,ab OR tofranil:ti,ab OR janimine:ti,ab OR praminil:ti,ab OR imiprex:ti,ab OR elepsin:ti,ab OR lomont:ti,ab OR gamanil:ti,ab OR deanxit:ti,ab OR dixeran:ti,ab OR melixeran:ti,ab OR trausabun:ti,ab OR timaxel:ti,ab OR pamelor:ti,ab OR aventyl:ti,ab OR norpress:ti,ab OR agedal:ti,ab OR elronon:ti,ab OR nogedal:ti,ab OR azafen:ti,ab OR azaphen:ti,ab OR vagran:ti,ab OR vivactil:ti,ab OR kevopril:ti,ab OR kinupril:ti,ab OR adeprim:ti,ab OR quinuprine:ti,ab

#	Search Terms
	OR survector:ti,ab OR maneon:ti,ab OR directim:ti,ab OR prondol:ti,ab OR galatur:ti,ab OR tetran:ti,ab OR insidon:ti,ab OR pramolan:ti,ab OR ensidon:ti,ab OR oprimol:ti,ab OR stablon:ti,ab OR coaxil:ti,ab OR tatinol:ti,ab OR surmontil:ti,ab OR asendin:ti,ab OR deprilept:ti,ab OR ludiomil:ti,ab OR psymion:ti,ab OR mazanor:ti,ab OR sanorex:ti,ab OR tecipul:ti,ab OR marplan:ti,ab OR aurorix:ti,ab OR manerix:ti,ab OR nardil:ti,ab OR eldepryl:ti,ab OR zelapar:ti,ab OR emsam:ti,ab OR solian:ti,ab OR clozaril:ti,ab OR nipolept:ti,ab OR remergil:ti,ab OR remeron:ti,ab OR zispin:ti,ab OR remergon:ti,ab OR rexer:ti,ab OR promyrtil:ti,ab OR norset:ti,ab OR 'remeron soltab':ti,ab OR '6 azamianserin':ti,ab OR mepirzepine:ti,ab OR 'org 3770':ti,ab OR brintellix:ti,ab OR valdoxan:ti,ab OR melitor:ti,ab OR thymanax:ti,ab OR deptran:ti,ab OR sinequan:ti,ab OR edronax:ti,ab OR prolift:ti,ab OR 'opc 34712':ti,ab OR viibryd:ti,ab OR serzone:ti,ab
27	#20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26
28	#9 AND #27
29	#28 AND [2013-2019]/py
30	#29 AND #16
31	#19 OR #30
32	('animal'/exp OR 'nonhuman'/exp OR 'animal experiment'/exp) NOT 'human'/exp
33	#31 NOT #32
34	#33 AND [english]/lim
35	#34 AND [medline]/lim
36	#34 NOT #35
37	#36 NOT ('case report'/de OR 'human tissue'/de OR 'nonhuman'/de OR 'practice guideline'/de OR 'questionnaire'/de OR 'chapter'/it OR 'conference review'/it OR 'editorial'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)
Date of search: December 5, 2018	

Figure A1. PRISMA Flow Chart Showing Results of Literature Search for TRD Treatment Options



Appendix B. Previous Systematic Reviews and Technology Assessments

Canadian Agency for Drugs and Technologies in Health (CADTH). 2017
Ketamine for Treatment-Resistant Depression or Post-Traumatic Stress Disorder in Various Settings

The Canadian Agency for Drugs and Technologies in Health (CADTH) evaluated the clinical effectiveness and safety, as well as evidence-based guidelines on the off-label use of ketamine for the treatment of treatment-resistant depression (TRD) and post-traumatic stress disorder (PTSD) in various settings. Three systematic reviews, five primary studies, and two evidence-based guidelines were included in the report. It was found that ketamine was effective in rapidly reducing symptom severity, suicidality, and fatigue in TRD, as well as in diminishing symptom severity in PTSD patients. Albeit being transient, ketamine's antidepressant effects were found to be comparable or superior to other pharmacological or somatic interventions (e.g., SSRIs or ECT) for TRD. Furthermore, use of ketamine did not compromise neurocognitive functioning, and serious adverse events were rare in the trials. Regardless of the cited clinical benefits, the guidelines included in this review recommended restricting access to ketamine to research settings. The authors noted the need for better quality studies, with larger sample sizes, longer follow-up, and repeated dosing to make more informed clinical guideline recommendations for the use of ketamine in patients with TRD.

Papadimitropoulou K, Vossen C, Karabis A, Donatti C, Kubitz N. Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis. *Current Medical Research and Opinion.* 2017;33(4):701-711.

The investigators performed a network meta-analysis (NMA) in order to indirectly compare and rank the relative efficacy and safety of pharmacological and somatic interventions for the treatment of TRD. The NMA included 31 randomized controlled trials (RCTs) that compared TRD treatments in adult patients. Results of the evidence synthesis showed the antidepressant effects of ketamine to be superior to pharmacological or somatic treatments at two weeks of treatment by showing higher response rates and a faster reduction in symptom severity. Ketamine data was not available for later timepoints (e.g., four, six, and eight weeks). In order to assess its long-term antidepressant efficacy and safety, additional ketamine studies are needed. With the exception of high dose quetiapine augmentation and risperidone augmentation, which were found to show superior outcomes, efficacy results at four, six, and eight weeks showed no clear distinction among the treatments that were investigated. Networks for response and remission rate outcomes were small at most timepoints. Lamotrigine augmentation, with a profile comparable to placebo/sham, was found to be the best tolerated treatment. It was concluded that long-term data that would allow for a comparative assessment of long-term efficacy is lacking and further studies are needed.

Appendix C. Ongoing Studies

Title/ Trial Sponsor	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
Esketamine					
An Open-label Long-term Extension Safety Study of Intranasal Esketamine in Treatment-resistant Depression (SUSTAIN-3) NCT02782104 Sponsor: Janssen Research & Development, LLC	Phase III Long-term, open-label extension safety study, single group assignment Estimated enrollment: 1150 (recruiting)	Experimental: **Esketamine** – In open-label induction phase (4 weeks) participants will self-administer intranasal esketamine twice weekly as a flexible dose regimen – In optimization/ maintenance phase (variable length) participants will self-administer intranasal esketamine once weekly (with option to individualize dosing frequency)	Inclusion criteria: ≥18 years of age, medically stable Based on prior study participant is entering from: – TRANSFORM-1 or TRANSFORM-2: Participant has completed induction phase and the 2 weeks follow up phase visit; or participant completed the induction phase and was a responder – SUSTAIN-1: Participant relapsed during the maintenance phase; participant completed study – SUSTAIN-2: Participant completed study – TRANSFORM-3: Participant was in the induction phase of the study at the time enrollment into the SUSTAIN-2 study was closed Exclusion criteria: – Since the last study visit in the participant's prior study, participant has suicidal ideation with intent to act, or suicidal behavior	*Time Frame: baseline of each dosing session (pre-dose) up to the last post-dose measurement from the start of induction phase to end of optimization/maintenance phase (approx. 5 years 3 months)* Primary Outcome Measures: – Number of participants with TEAEs – Change from baseline in SBP and DBP; HR – Change from baseline in MOAAS score (*1-hour post-dose for duration of the study*) Secondary Outcome Measures: – Change from baseline in participant-reported depressive symptoms and CGI-S score	August 2019

Title/ Trial Sponsor	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
			– Evidence of previous drug use on the day of the first intranasal treatment session; or has taken any prohibited therapies		
A Randomized, Double-blind, Multicenter Active-controlled Study to Evaluate the Efficacy, Pharmacokinetics, Safety and Tolerability of Flexible Doses of Intranasal Esketamine Plus an Oral Antidepressant in Adult Subjects With Treatment-resistant Depression NCT03434041 Sponsor: Janssen Research & Development, LLC	Phase III Randomized, double-blind trial, parallel assignment Estimated enrollment: 234 (recruiting)	Experimental: **Esketamine + AD** – In double-blind treatment phase (4 weeks) participants will self-administer esketamine (flexible dosing) intranasally twice weekly – Participants will initiate a new, open-label oral antidepressant during double-blind treatment phase Active Comparator: **Placebo + AD** – Participants will self-administer matching placebo intranasally twice weekly during double-blind treatment phase – Participants will initiate a new, open-label oral antidepressant during double-blind treatment phase	Inclusion Criteria: – 18-64 years, medically stable – DSM-5 diagnosis for recurrent or single-episode MDD, without psychotic features – Non-response to 1-5 oral antidepressant treatments in current episode (if duration >2 years) – MADRS score ≥28 Exclusion Criteria: – Previous non-response to esketamine, ketamine, or all oral antidepressant options available; treatment with ECT, VNS, or DBS in current MDD episode – Current or prior DSM-5 diagnosis of a psychotic disorder; MDD with psychotic features; bipolar or related disorders; OCD (current episode only); intellectual disability; autism spectrum disorder; borderline and connected personality disorders – Homicidal ideation, with some	Primary Outcome: – Change from baseline in MADRS score *[Time Frame: Baseline and end of double-blind treatment phase]* Secondary Outcomes: – Percentage of participants with onset of clinical response *[Time Frame: Day 2 through end of double-blind treatment phase]* – Change from baseline in SDS and CGI-S scale *[Time Frame: Baseline and end of double-blind treatment phase]* – Percentage of responders and participants in remission *[Time Frame: At end of double-blind treatment phase]* – Percentage of participants with sustained remission *[Time Frame: Up to 8 weeks]*	April 2021

Title/ Trial Sponsor	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
			intent to act within 6 months prior to screening		
A Randomized, Double-blind, Multicenter, Placebo-controlled Study to Evaluate the Efficacy, Safety and Tolerability of Fixed Doses of Intranasal Esketamine in Japanese Subjects with Treatment Resistant Depression NCT02918318 Sponsor: Janssen Pharmaceutical K.K	Phase II Randomized, double-blind trial, parallel assignment Estimated enrollment: 183 (recruiting)	Fixed dosing (28mg esketamine per spray); 4-week induction phase (double-blind or open-label); Post-treatment phase: 193 days Experimental: – **Esketamine, 28 mg**: Participant will receive 1 spray of esketamine to each nostril at 0 minutes and placebo at 5 and 10 minutes – **Esketamine, 56 mg**: Participant will receive 1 spray of esketamine to each nostril at 0 and 5 minutes, and placebo at 10 minutes – **Esketamine, 84 mg**: Participant will receive 1 spray of Esketamine to each nostril at 0, 5, and 10 minutes Comparator: **Placebo** Participant will receive 1 spray of placebo to each nostril at 0, 5, and 10 minutes	Inclusion Criteria: – 20-64 years, medically stable – DSM-5 diagnosis of single-episode (persistent; duration ≥2 years) or recurrent MDD, without psychotic features – MADRS score ≥28 and antidepressant treatment non-response in current episode Exclusion Criteria: Participant has received VNS or DBS treatment in the current episode Previous treatment with esketamine or ketamine Homicidal or suicidal ideation or intent to act within 6 months History of SUD according to DSM-5 criteria within 6 months of screening phase Current or history of seizure disorder	Primary Outcomes: – Change in MADRS Score *[Time Frame: Baseline up to end of the double-blind induction phase (day 28)]* Secondary Outcomes: *[Time Frame: From baseline to Day 28]* Percentage of responders and remitters Change in MADRS; CGI-S score; SDS in double-blind induction phase Proportion of responders and remitters Time to relapse in patients who responded/remitted at end of double-blind induction phase [Time Frame: Performed weekly through week 24 or relapse in post-treatment phase] Change in SDS score in post-treatment phase [Time Frame: Baseline up to end of post-treatment phase (day 193)]	February 2019
A Study of Esketamine Nasal Spray Plus a New Standard-of-care Oral	Phase III Randomized, double-blind	Experimental: Esketamine + AD – **Esketamine, 28 mg**: Initial dose for elderly participants (65-74 years), then uptitrated to 56 mg on day 4	Inclusion Criteria: – 18-74 years, medically stable – DSM-5 diagnosis of single-episode	Primary Outcomes: Percentage of participants with remission (MADRS score ≤10) at the end of week 8	July 2021

Title/ Trial Sponsor	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
Antidepressant or Placebo Nasal Spray Plus a New Standard-of-care Oral Antidepressant in Adult and Elderly Participants With Treatment-resistant Depression NCT03852160 Sponsor: Janssen-Cilag International NV	trial, parallel assignment Estimated Enrollment: 580 (not yet recruiting)	– **Esketamine, 56 mg**: Initial dose for participants ≤64 years. Dose may be increased – **Esketamine, 84 mg**: maximum uptitrated esketamine dose Active comparator: Matching placebo + AD *Dosing regimen*: – Intranasal, twice-weekly with a flexible dose regimen from Day 1 until Day 28 (Week 4) – Intranasal, once weekly from week 5 to week 8 – Intranasal, once weekly or once every other week from Week 9 to Week 32	or recurrent MDD, without psychotic features; non-response to 2-6 antidepressant treatments in current episode – IDS-C30 score ≥34 Exclusion Criteria: – Previous non-response to (a) esketamine or ketamine; (b) all antidepressant classes available in the study or augmentation/combination therapy in the current episode; (c) ECT treatment in current episode – Received VNS or DBS in current episode – Current or prior DSM-5 diagnosis of a psychotic disorder or MDD with psychotic features; bipolar or related disorders; intellectual disability; autism spectrum disorder; borderline personality disorder; antisocial personality disorder – Homicidal ideation or intent; suicidal ideation with some intent to act within 1 month prior to screening – History of SUD or severe alcohol use disorder within 6 months of study screening	Secondary Outcomes: Percentage of participants with remission at week 8 w/o relapse until week 32 Change in MADRS score from baseline at week 4 *[Time Frame: Baseline, up to week 32]* Change from baseline in MADRS; CGI-S; SDS Medical resource utilization; number of participants with TEAEs Suicidal ideation and behavior	

Title/ Trial Sponsor	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date

Source: www.ClinicalTrials.gov (NOTE: studies listed on site include both clinical trials and observational studies). CGI-S Scale: Clinical Global Impression Severity Scale, DBP: Diastolic Blood Pressure, DBS: Deep Brain Stimulation, DSM-V: Diagnostic and Statistical Manual (5th edition), ECT: Electroconvulsive Therapy, HR: Heart Rate, IDS: Inventory of Depressive Symptomatology, MADRS: Montgomery-Åsberg Depression Rating Scale, MDD: Major Depressive Disorder, MOAAS Scale: Modified Observer's Assessment of Alertness/Sedation Scale, OCD: Obsessive Compulsive Disorder, SBP: Systolic Blood Pressure, SDS: Sheehan Disability Scale, SUD: Substance Use Disorder, TEAE: Treatment Emergent Adverse Events, VNS: Vagus Nerve Stimulation.

Appendix D. Comparative Clinical Effectiveness Supplemental Information

We performed screening at both the abstract and full-text level. Three investigators screened abstracts identified through electronic searches according to the inclusion and exclusion criteria described earlier. We did not exclude any study at abstract-level screening due to insufficient information. For example, an abstract that did not report an outcome of interest would be accepted for further review in full text. We retrieved the citations that were accepted during abstract-level screening for full text appraisal. Three investigators reviewed full papers and provided justification for exclusion of each excluded study.

We also included FDA documents related to esketamine. These included the manufacturer's submission to the agency, internal FDA review documents, and the transcript of Advisory Committee deliberations and discussions. All literature that did not undergo a formal peer review process is described separately.

We used criteria published by the US Preventive Services Task Force (USPSTF) to assess the quality of RCTs and comparative cohort studies, using the categories "good," "fair," or "poor" (see Appendix Table D1 and D7)[110] Guidance for quality ratings using these criteria is presented below, as is a description of any modifications we made to these ratings specific to the purposes of this review.

Good: *Meets all criteria: Comparable groups are assembled initially and maintained throughout the study; reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention is paid to confounders in analysis. In addition, intention to treat analysis is used for RCTs.*

Fair: *Studies were graded "fair" if any or all of the following problems occur, without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are addressed. Intention to treat analysis is done for RCTs.*

Poor: *Studies were graded "poor" if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat analysis is lacking.*

Note that case series are not considered under this rating system – because of the lack of comparator, these are generally considered to be of poor quality.

Table D1. Quality of Included RCTs of Esketamine*

Trial	Comparable Groups	Non-Differential Follow-up	Patient/ Investigator Blinding	Clear Definition of Intervention	Clear Definition of Outcomes	Selective Outcome Reporting	Measurements Valid	ITT Analysis	Approach to Missing Data	USPSTF Rating
TRANSFORM-1[50]	Yes	No	Yes	Yes	Yes	*	Yes	mITT	MMRM	*
TRANSFORM-2[51]	Yes	Yes	Yes	Yes	Yes	*	Yes	mITT	MMRM	*
TRANSFORM-3[52]	Yes	Yes	Yes	Yes	Yes	*	Yes	mITT	MMRM	*
SUSTAIN-1[53]	Yes	Yes	Yes	Yes	Yes	*	Yes	mITT	NR	*

ITT: intention to treat, mITT: modified intention to treat, MMRM: mixed-effects model using repeated measures, NR: not reported, USPSTF: US Preventive Services Task Force

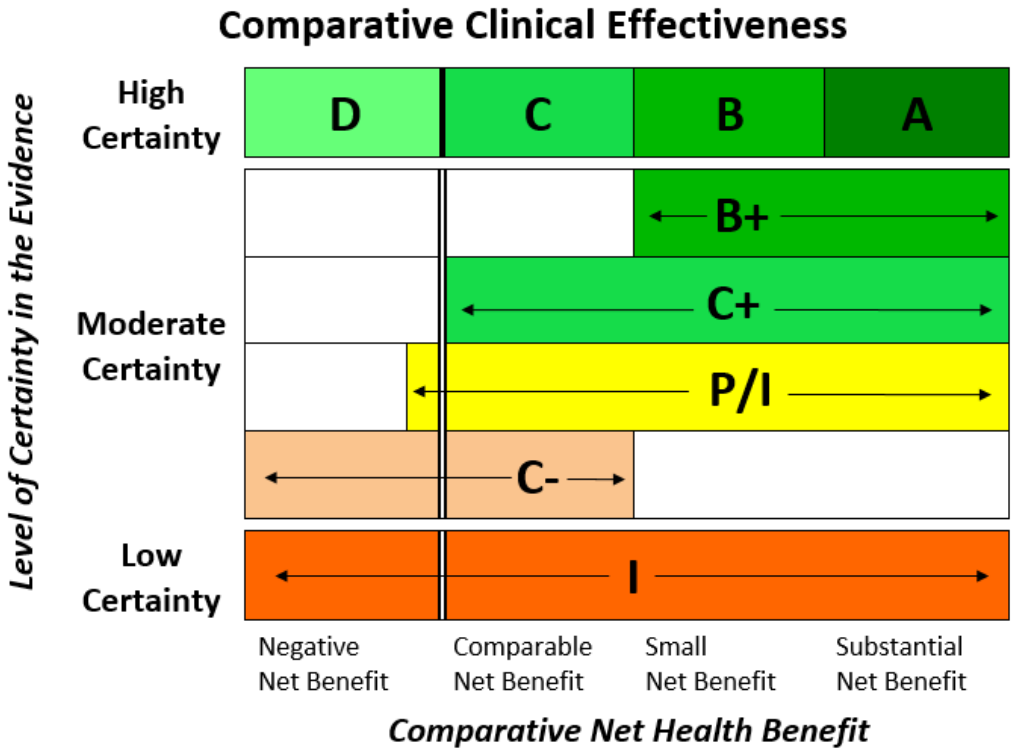
*The data for the esketamine trials were only available in grey literature. Due to this, we did not assign an overall quality rating for the trials and were not able to assess selective outcome reporting. We will assign an overall quality rating and update quality categories where necessary upon publication of peer-reviewed results.

ICER Evidence Rating

We used the ICER Evidence Rating Matrix (see Figure D1) to evaluate the evidence for a variety of outcomes. The evidence rating reflects a joint judgment of two critical components:

a) The **magnitude** of the difference between a therapeutic agent and its comparator in "net health benefit" – the balance between clinical benefits and risks and/or adverse effects AND
b) The level of **certainty** in the best point estimate of net health benefit.[49]

Figure D1. ICER Evidence Rating Matrix



***A = "Superior"** - High certainty of a substantial (moderate-large) net health benefit*
***B = "Incremental"** - High certainty of a small net health benefit*
***C = "Comparable"**- High certainty of a comparable net health benefit*
***D = "Negative"**- High certainty of an inferior net health benefit*
***B+ = "Incremental or Better"** - Moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit*
***C+ = "Comparable or Better"** - Moderate certainty of a comparable, small, or substantial net health benefit, with high certainty of at least a comparable net health benefit*
***P/I = "Promising but Inconclusive"** - Moderate certainty of a comparable, small, or substantial net health benefit, and a small (but nonzero) likelihood of a negative net health benefit*
***C- = "Comparable or Inferior"** - Moderate certainty that the point estimate for comparative net health benefit is either comparable or inferior*
***I = "Insufficient"** - Any situation in which the level of certainty in the evidence is low*

Table D2. Study Design of Phase III Trials of Esketamine

Trial	Key Study Phases	Study Arms & Dosing	Key Inclusion Criteria	Key Exclusion Criteria
TRANSFORM-1[50] **Phase III Multicenter, Global**	4-week prospective observational phase; 4-week randomized double-blind induction phase; 24-week follow-up or patients enter SUSTAIN-1	1) Esketamine 56 mg + AD (n=115) 2) Esketamine 86 mg + AD (n=114) 3) Placebo + AD (n=113) Intranasal esketamine and placebo were administered twice weekly. Newly-initiated, open-label oral AD (duloxetine, escitalopram, sertraline, venlafaxine extended release) was chosen by investigators and administered daily.	– 18-64 years – DSM-5 criteria for recurrent or single episode (≥2 years) MDD without psychotic features – MADRS≥28; ICD-C30≥34 – Failed 1-5 ADs in current episode at screening (≤25% improvement assessed by MGH-ATRQ) and prospectively failed AD during observational phase (≤25% improvement in MADRS and MADRS≥28 at weeks 2 and 4)	– Suicidal ideation with intent to act in prior 6 months – Suicidal behavior in prior year – Bipolar disorder or other current or prior DSM-5 psychotic disorder – Failed ECT, all SSRI/SNRI options offered in trial, ketamine, or esketamine in current episode – History of moderate-to-severe substance use disorder in prior 6 months
TRANSFORM-2[51] **Phase III Multicenter, US & Europe**	4-week prospective observational phase; 4-week randomized double-blind induction phase; 24-week follow-up or patients enter SUSTAIN-1	1) Esketamine (flexible: 56 or 84 mg) + AD (n=114) 2) Placebo + AD (n=109) Intranasal esketamine and placebo were administered twice weekly. Patients started with 56 mg on day 1. Dose could be flexibly increased to 84 mg or stay at 56 mg on days 4, 8, 11, or 15 (after which the dose remained stable). Oral AD was administered in same manner as in TRANSFORM-1.	Same criteria as TRANSFORM-1	Same criteria as TRANSFORM-1
TRANSFORM-3[52] **Phase III Multicenter, Global**	4-week prospective observational phase; 4-week randomized double-blind induction phase; 2-week follow-up or patients enter SUSTAIN-2	1) Esketamine (flexible: 28, 56, or 84 mg) + AD (n=72) 2) Placebo + AD (n=65) Intranasal esketamine and placebo were administered twice weekly. Patients started with 28 mg on day 1 and could flexibly titrate to 56 or 84	– ≥65 years – DSM-5 criteria for recurrent or single episode (≥2 years) MDD without psychotic features – MADRS≥24; ICD-C30≥31 – Failed 1-8 ADs in current episode at screening (≤25% improvement assessed by MGH-ATRQ geriatric version) and	Same criteria as TRANSFORM-1

Trial	Key Study Phases	Study Arms & Dosing	Key Inclusion Criteria	Key Exclusion Criteria
		mg at subsequent visits. Oral AD was administered in same manner as in TRANSFORM-1.	prospectively failed AD during observational phase (≤25% improvement in MADRS and MADRS≥24 at weeks 2 and 4)	
SUSTAIN-1[53] **Phase III Multicenter, Global**	12-week optimization phase; Up to 48-week randomized double-blind withdrawal phase;[1] 2-week follow-up [1]Patients who achieved stable remission (MADRS≤12 for ≥3 out of last 4 weeks) or stable response (≥50% reduction in MADRS in last 2 weeks) during the optimization phase were separately randomized to receive placebo or esketamine for up to 48 weeks until relapse.	Stable Remission: 1) Esketamine (flexible: 56 or 84 mg) + AD (n=90) 2) Placebo + AD (n=86) Stable Response: 1) Esketamine (flexible: 56 or 84 mg) + AD (n=62) 2) Placebo + AD (n=59) Intranasal esketamine and placebo were administered once weekly or every other week, depending on the severity of depressive symptoms. Patients continued the same dose of esketamine and oral AD they received during the optimization phase.	– Patients either enrolled directly or rolled over from TRANSFORM-1 or -2 – Patients who responded to treatment (≥50% reduction in MADRS) during 4-week induction phase entered the optimization phase – Direct-entry inclusion criteria is the same as TRANSFORM-1	Same criteria as TRANSFORM-1
SUSTAIN-2[54] **Phase III Multicenter, Global**	48-week optimization/ maintenance phase;[1] 4-week follow up [1]This phase was terminated when the predefined exposure criteria were met: ≥300 patients reached 6 months exposure and ≥100 patients reached 12 months exposure to esketamine.	1) Esketamine (flexible: 28, 56, or 84 mg) + AD (n=603) Patients received intranasal esketamine once weekly for 4 weeks and then individualized to once weekly or every other week for the remainder of the optimization/maintenance phase.	– Patients either enrolled directly or rolled over from TRANSFORM-3. Patients who responded (≥50% reduction in MADRS) during 4-week induction phase entered the optimization/maintenance phase. Direct-entry inclusion criteria: – ≥18 years – DSM-5 criteria for recurrent or single episode (≥2 years) MDD without psychotic features – MADRS≥22	Same criteria as TRANSFORM-1

Trial	Key Study Phases	Study Arms & Dosing	Key Inclusion Criteria	Key Exclusion Criteria
			– Failed ≥2 ADs in current episode assessed by MGH-ATRQ	

AD: Antidepressant, DSM-5: Diagnostic and Statistical Manual of Mental Disorders (5th Edition), ECT: Electroconvulsive Therapy, ICD-C30: Inventory of Depressive Symptomatology-Clinician rated (30-item), MADRS: Montgomery–Åsberg Depression Rating Scale, MDD: Major Depressive Disorder, MGH-ATRQ: Massachusetts General Hospital – Antidepressant Treatment Response Questionnaire, n=number of patients, SNRI: Serotonin and norepinephrine reuptake inhibitor, SSRI: Selective serotonin reuptake inhibitors

Table D3. Key Baseline Characteristics for Phase III Trials of Esketamine

Trial	Arm	N	Age (y), Mean (SD)	Age of MDD Diagnosis, Mean (SD)	Duration of Current Episode (y), Mean (SD)	≥3 Failed ADs at Baseline, n (%)	MADRS, Mean (SD)	PHQ-9, Mean (SD)	SDS, Mean (SD)	CGI-S, Mean (SD)	SNRI; SSRI , n (%)*
TRANSFORM-1 [50,84]	ESK 56 mg + AD	115	46.4 (11.2)	30.3 (12.3)	3.9 (5.3)	34 (30.1)	37.4 (4.8)	20.3 (4.1)	24.0 (4.1)	5.1 (0.7)	65 (56.5); 50 (43.5)
	ESK 84 mg + AD	114	45.7 (11.1)	32.1 (12.9)	4.1 (6.3)	55 (48.2)	37.8 (5.6)	20.7 (3.6)	24.7 (4.6)	5.1 (0.7)	67 (58.8); 47 (41.2)
	PBO + AD	113	46.8 (11.4)	31.8 (12.4)	3.7 (5.1)	46 (40.7)	37.5 (6.2)	20.8 (3.7)	24.4 (3.9)	5.1 (0.7)	64 (56.6); 49 (43.4)
TRANSFORM-2 [51,84]	ESK# + AD	114	44.9 (12.6)	32.1 (12.5)	2.1 (2.4)	36 (31.6)	37.0 (5.7)	20.2 (3.6)	24.0 (4.1)	5.1 (0.7)	77 (67.5); 37 (32.5)
	PBO + AD	109	46.4 (11.1)	35.3 (13.0)	2.3 (3.6)	37 (39.9)	37.3 (5.7)	20.4 (3.7)	24.2 (4.4)	5.1 (0.7)	75 (68.8); 34 (31.2)
TRANSFORM-3 [52,84]	ESK# + AD	72	70.6 (4.8)	42.6 (16.2)	3.1 (5.3)	26 (36.1)	35.5 (5.9)	17.6 (5.0)	21.8 (5.9)	NR	31 (41.3); 41 (56.9)
	PBO + AD	65	69.4 (4.2)	43.7 (16.3)	5.3 (7.6)	27 (41.5)	34.8 (6.4)	17.4 (6.3)	22.9 (4.7)	NR	30 (46.2); 35 (53.8)
SUSTAIN-1 [53,84]	All enrolled patients‡	705	46.1 (11.1)	32.7 (11.7)	NR	NR	37.9 (5.5)	19.9 (4.2)	23.8 (4.4)	NR	440 (63); 259 (37)†
	Stable Remission‡: ESK# + AD	90	45.4 (12.1)	32.5 (11.4)	2.2 (3.3)	27 (30.0)	37.4 (5.2)	19.2 (4.2)	23.5 (3.4)	NR	62 (68.9); 28 (31.1)
	Stable Remission‡: PBO + AD	86	46.2 (11.2)	33.4 (11.4)	2.1 (2.8)	28 (32.6)	37.6 (4.6)	19.8 (3.4)	23.8 (4.0)	NR	58 (67.4); 28 (32.6)
	Stable Response‡: ESK# + AD	62	47.2 (11.0)	36.2 (13.3)	2.3 (3.7)	27 (43.5)	40.1 (5.6)	20.5 (4.1)	24.8 (3.6)	NR	35 (56.5); 27 (43.5)
	Stable Response‡: PBO + AD	59	46.7 (9.8)	34.0 (10.5)	2.7 (4.9)	27 (45.8)	38.9 (4.9)	20.4 (4.2)	24.0 (3.7)	NR	36 (61.0); 23 (39.0)
SUSTAIN-2 [54,84]	All enrolled patients‡	802	52.2 (13.7)	35.7 (13.8)	3.1 (5.0)	320 (39.9)	31.4 (5.4)	17.3 (5.0)	22.2 (5.4)	NR	407 (51); 394 (49)§

AD: antidepressant, CGI-S: Clinical Global Impression-Severity, ESK: esketamine, MADRS: Montgomery–Åsberg Depression Rating Scale, MDD: Major depressive disorder, n: number of individuals, N: n at randomization, PBO: placebo, PHQ-9: Patient Health Questionnaire, SD: standard deviation, SDS: Sheehan Disability Scale, SNRI: Serotonin and norepinephrine reuptake inhibitor, SSRI: Selective serotonin reuptake inhibitors, y: years.

*Percent of patients receiving class of AD during the trial as assigned by the investigator at randomization; †Data available for 699 patients; ‡Characteristics at the beginning of the induction phase; §One patient did not receive oral AD; #Flexible dosing.

Table D4. Key Efficacy Outcomes from the Short-Term Trials of Esketamine at Week Four

Trial	Arm	Change in MADRS			Response*	Remission†	Change in PHQ-9			Change in SDS			Onset of Sustained Clinical Response‡	
		N	Mean Change (SD)	Difference in LSM (95% CI); p-value	n (%)	n (%)	N	Mean Change (SD)	Difference in LSM (95% CI); p-value	N	Mean Change (SD)	Difference in LSM (95% CI); p-value	n (%)	OR (95% CI); P-Value
TRANSFORM-1[50,84]	ESK 56 mg + AD	111	-19.0 (13.9)	-4.1 (-7.7, -0.5); 0.0114§	60 (54.1)	40 (36.0)	110	-11.0 (8.1)	-2.3 (-4.3, -0.3); NA#	88	-11.0 (9.3)	-2.5 (-5.3, 0.20); NA#	12 (10.4)	6.5 (1.4, 60.5); NA#
	ESK 84 mg + AD	98	-18.8 (14.1)	-3.2 (-6.9, 0.5); NS	52 (53.1)	38 (38.8)	99	-11.7 (7.7)	-2.2 (-4.3, -0.2); NA#	87	-11.1 (10.0)	-2.2 (-4.9, 0.5); NA#	10 (8.8)	5.3 (1.1, 50.9); NA#
	PBO + AD	108	-14.8 (15.1)	---	42 (38.9)	33 (30.6)	108	-9.1 (8.4)	---	90	-8.4 (9.7)	---	2 (1.8)	---
TRANSFORM-2[51]	ESK¤ + AD	101	-21.4 (12.3)	-4.0 (-7.3, -0.6); 0.020	70 (69.3)	53 (52.5)	104	-13.0 (6.4)	-2.4 (-4.2, -0.7); 0.006	86	-13.6 (8.3)	-4.0 (-6.3, -1.6); <0.001	9 (7.9)	1.79; (0.6, 5.7); 0.321
	PBO + AD	100	-17.0 (13.9)	---	52 (52.0)	31 (31.0)	100	-10.2 (7.8)	---	85	-9.4 (8.4)	---	5 (4.6)	---
TRANSFORM-3[52,84]	ESK¤ + AD	63	-10.0 (12.7)	-3.6 (-7.2, 0.1); NS	17 (27.0)	11 (17.5)	NR	NR	-2.8** (-5.1, -0.5); NR	NR	NR	-4.6** (-8.3, -1.0); NR	NR	
	PBO + AD	60	-6.3 (8.9)	---	8 (13.3)	4 (6.7)	NR	NR	---	NR	NR	---		

CI: Confidence interval, AD: antidepressant, ESK: esketamine, LSM: least square mean, MADRS: Montgomery–Åsberg Depression Rating Scale, n: number of individuals, N: n analyzed, NA: not applicable, NS: not significant, OR: odds ratio, PBO: placebo, PHQ-9: Patient Health Questionnaire, SD: standard deviation, SDS: Sheehan Disability Scale. *≥50% reduction in MADRS from baseline to week 4; †MADRS≤12 at week 4; ‡50% reduction in MADRS by day 2 maintained until day 28; §P-value from exploratory analysis reported in FDA Briefing Document[83]; #Per the predefined statistical plan, the statistical significance of secondary endpoints was not formally evaluated if the primary endpoint did not meet statistical significance; ¤Flexible dosing; **Digitized estimate

Table D5. Key Efficacy Outcomes from the Randomized Withdrawal Study and the Long-term Safety Study of Esketamine at Endpoint*

Trial	Arm	N	Relapse†	Time to Relapse (days)		Response‡	Remission§	Change in MADRS
			n (%)	Median (95% CI)	HR (95%CI); P-Value	n (%)	n (%)	Mean Change (SD)
SUSTAIN-1[53]	Stable Remission: ESK# + AD	90	24 (26.7)	NE¤	0.49 (0.29, 0.84); 0.003	Not measured		
	Stable Remission: PBO + AD	86	39 (45.3)	273 (97.0, NE¤)	---			
	Stable Response: ESK# + AD	62	16 (25.8)	635 (264, 635)	0.30 (0.16, 0.55) <0.001			
	Stable Response: PBO + AD	59	34 (57.6)	88 (46, 196)	---			
SUSTAIN-2[54]	ESK# + AD	603	Not measured			461 (76.5)	351 (58.2)	0.3 (8.1)**

CI: Confidence Interval, AD: antidepressant, ESK: esketamine, HR: hazard ratio, MADRS: Montgomery–Åsberg Depression Rating Scale, n: number of individuals, N: n analyzed, NE: not estimable, PBO: placebo, SD: standard deviation

*Timepoint at which outcomes were measured varies. In SUSTAIN-1, patients were followed up to 48 weeks until relapse. In SUSTAIN-2, patients were followed for up to one year; the study was terminated early when the predefined exposure criteria were met. †MADRS≥22 for two consecutive visits separated by 5-15 days or hospitalization for any event suggestive of relapse (e.g., worsening depression, suicide attempt, suicide prevention, completed suicide); ‡≥50% reduction in MADRS from baseline at endpoint; §MADRS≤12 at endpoint; #Flexible dosing; ¤50% relapse rate not reached based on Kaplan Meier estimates; **Change from baseline of optimization/maintenance phase to end of optimization/maintenance phase

Table D6. Treatment-Emergent Adverse Events in Phase III Trials of Esketamine

Trial	TRANSFORM-1[50,84]			TRANSFORM-2[51,84]		TRANSFORM-1 & -2 pooled[84]		TRANSFORM-3[84]		SUSTAIN-1[84]		SUSTAIN-2[84]		
Week	4			4		4		4		≥48		4	≥48	≥52
Arm	ESK 56 mg + AD	ESK 84 mg + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	IND Phase: ESK + AD	OP/ MAINT Phase: ESK+ AD	Both Phases: ESK + AD
N	115	116	113	115	109	346	222	72	65	152	145	779	603	802
Any TEAE, %	87.0	88.8	68.1	85.2	60.6	87.0	64.4	70.8	60.0	82.2	45.5	83.8	85.6	90.1
Serious TEAE, %	1.7	0	0	0.9	0.9	0.9	0.5	4.2	3.1	2.6	0.7	2.2	6.3	6.9
D/C Due to TEAE, %	0.9	6.0	---	7.0	---	4.6	1.4	5.6	3.1	2.6	2.1	6.8	3.8	9.5
Death, %	0	0	0	0.9	0	0.3	0	0	0	0	0	0	0.3	0.2
Dizziness, %	27.8	22.4	8.8	20.9	4.6	23.7	6.8	22.2	7.7	20.4	4.8	29.3	22.4	32.9
Dissociation, %	26.1	27.6	3.5	26.1	3.7	26.6	3.6	12.5	1.5	23.0	0	23.4	18.7	27.6
Headache, %	20.0	20.7	16.8	20.0	17.4	20.2	17.1	12.5	3.1	17.8	9.7	17.6	19.1	25.1
Nausea, %	27.0	31.9	10.6	26.1	6.4	28.3	8.6	18.1	4.6	16.4	0.7	20.2	13.9	25.1
Somnolence, %	20.9	18.1	11.5	13.0	6.4	17.3	9.0	---	---	21.1	2.1	12.1	14.1	16.7
Dysgeusia, %	14.8	17.2	15.0	24.3	11.9	18.8	13.5	5.6	4.6	27.0	6.9	9.9	9.0	11.8
Vertigo, %	20.9	20.7	1.8	26.1	2.8	22.5	2.3	11.1	3.1	25.0	5.5	8.7	7.1	11.0
Hypoesthesia, %	12.2	13.8	1.8	7.0	0.9	11.0	1.4	5.6	1.5	5.9	0	10.1	6.6	11.8
Vomiting, %	6.1	12.1	1.8	9.6	1.8	9.2	1.8	6.9	1.5	6.6	0.7	7.2	7.5	10.8
BP Increase, %	7.8	9.5	4.4	9.6	0	9.0	2.3	12.5	4.6	6.6	3.4	6.8	7.8	9.5
Insomnia, %	8.7	6.9	9.7	9.6	4.6	8.4	7.2	5.6	4.6	---	---	5.3	5.8	8.1
Hypoesthesia Oral, %	13.9	10.3	1.8	7.8	0.9	10.7	1.4	6.9	0	13.2	0	8.1	---	9.1
Anxiety, %	8.7	7.8	6.2	10.4	4.6	9.0	5.4	2.8	7.7	7.9	4.1	6.5	---	9.0
Dizziness Postural, %	6.1	6.0	0	7.0	0.9	6.4	0.5	---	---	6.6	2.1	6.9	6.8	8.4
Paresthesia, %	16.5	9.5	2.7	11.3	0.9	12.4	1.8	5.6	3.1	7.2	0	5.9	---	7.2
Vision Blurred, %	7.0	7.8	0	12.2	2.8	9.0	1.4	---	---	15.8	0.7	6.3	---	7.5
Fatigue, %	10.4	6.9	4.4	4.3	5.5	7.2	5.0	12.5	7.7	---	---	5.1	---	7.9
Sedation, %	5.2	6.9	0.9	4.3	0.9	5.5	0.9	0	0	6.6	0.7	6.5	---	8.9
Diarrhea, %	7.0	4.3	2.7	8.7	9.2	6.6	5.9	---	---	---	----	---	6.5	7.5

Trial	TRANSFORM-1[50,84]			TRANSFORM-2[51,84]		TRANSFORM-1 & -2 pooled[84]		TRANSFORM-3[84]		SUSTAIN-1[84]		SUSTAIN-2[84]		
Week	4			4		4		4		≥48		4	≥48	≥52
Arm	ESK 56 mg + AD	ESK 84 mg + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	ESK + AD	PBO + AD	IND Phase: ESK + AD	OP/ MAINT Phase: ESK+ AD	Both Phases: ESK + AD
UTI %	---	---	---	---	---	---	---	8.3	1.5	---	---	---	---	8.1
Throat Irritation, %	4.3	7.8	3.5	7.8	4.6	6.6	4.1	---	---	5.3	0.7	---	---	---
Nasal Discomfort, %	3.5	4.3	6.2	7.0	1.8	4.9	4.1	---	---	7.2	2.8	---	---	---
Dry Mouth, %	4.3	4.3	3.5	7.8	2.8	5.5	3.2	---	---	---	---	---	---	---
Paresthesia Oral, %	7.8	0.9	1.8	7.8	0.9	5.5	1.4	---	---	5.3	0.7	---	---	---
Feeling Drunk, %	6.1	2.6	0	7.8	0.9	5.5	0.5	---	---	---	---	---	---	---

AD: antidepressant, BP: blood pressure, D/C: discontinuation, ESK: esketamine, IND: induction, N: number of patients analyzed, OP/MAINT: optimization/maintenance, PBO: placebo, TEAE: treatment-emergent adverse event, UTI: urinary tract infection.

Table D7. Key Characteristics and Outcomes of RCTs of Comparators (Ketamine, Olanzapine, rTMS, ECT)

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
Ketamine						
Singh 2016[55] Multicenter, US 4-week double-blind phase *Poor quality*	1) Placebo 2x/week + AD (n=17) 2) Ketamine 0.5 mg/kg 2x/week + AD (n=18) 3) Placebo 3x/week + AD (n=16) 4) Ketamine 0.5 mg/kg 3x/week + AD (n=17) All continued AD (most commonly used were fluoxetine, citalopram, and bupropion)	Failed ≥2 ADs in history, with ≥1 failure in current episode assessed by MGH-ATRQ	– Mean age: 43.9 years – Mean MADRS: 35.2 – Duration of current episode: NR – % with ≥3 AD failures in current episode: 15.0	<u>MADRS, mean change (SD)</u> At week 2: 1) -5.7 (10.2) 2) -18.4 (12), *p<0.001* vs 1 3) -3.1 (5.7) 4) -17.7 (7.3), *p<0.001* vs 3 At week 4: 1) -4.0 (9.1) 2) -21.2 (12.9) 3) -3.6 (6.6) 4) -21.1 (11.2) *p-values not reported*	<u>Response (MADRS), %</u> At week 2: 1) 15.4 2) 68.8, *p=0.005* vs 1 3) 6.3 4) 53.8, *p=0.004* vs 3 <u>Remission: MADRS≤10, %</u> At week 2: 1) 7.7 2) 37.5, *p=0.05* vs 1 3) 0 4) 23.1, *NS* vs 3	The most common TEAEs with incidence ≥20% and occurring more frequently in the ketamine arms were headache, anxiety, dissociation, nausea, and dizziness. There were two SAEs in patients receiving ketamine (anxiety and suicide attempt).
Olanzapine						
Shelton 2005[69] Multicenter, US & Canada 8-week double-blind phase *Good quality*	1) Olanzapine + Fluoxetine [OFC] (n=146) 2) Olanzapine + PBO [OLZ] (n=144) 3) Fluoxetine + PBO [FLX] (n=142) 4) Nortriptyline + PBO [NRT] (n=68) OFC and FLX arms initiated new AD; OLZ arm did not receive concomitant AD; NRT arm continued AD	≥1 SSRI failure in history + prospective nonresponse (≤30% decrease in MADRS) to nortriptyline	– Mean age: 42.4 years – Mean MADRS: 28.5 – Median duration of current episode: 1.0 year – No. of AD failures in current episode: NR	<u>MADRS, mean change (SE)</u> 1) -8.7 (0.7) *2)* -7.0 (0.7), *NS* vs 1 *3)* -8.5 (0.7), *NS* vs 1 *4)* -7.5 (1.0), *NS* vs 1	<u>Response (MADRS), %</u> 1) 27.5 2) 19.3 3) 28.9 4) 30.3 <u>Remission: MADRS≤8 at 2 consecutive visits, %</u> 1) 16.9 2) 12.9 3) 13.3 4) 18.2 *p NS among all groups for both outcomes. Pairwise*	More patients treated with OFC experienced ≥10% increase in weight compared to FLX (p=0.001) and NRT (p=0.02). Tremors occurred at a higher incidence in patients treated with OFC compared to FLX (p<0.001) and OLZ (p=0.053).

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
					p-values NR.	
Corya 2006[68] Multicenter, 16 countries (NR) 12-week double-blind phase *Fair quality*	1) Olanzapine + Fluoxetine [OFC] (four highest doses combined) (n=243) 2) Olanzapine + PBO [OLZ] (n=62) 3) Fluoxetine + PBO [FLX] (n=60) 4) Venlafaxine + PBO [VNL] (n=59) 5) OFC (1/5 mg; pseudo-placebo) (n=59) OFC and FLX arms initiated new AD; OLZ arm did not receive concomitant AD; VNL arm continued AD	1 SSRI failure in history + prospective nonresponse (≤30% decrease in MADRS) to venlafaxine	– Mean age: 45.7 years – Mean MADRS: 30.0 – Median duration of current episode: 0.5 years – No. of AD failures in current episode: NR	MADRS, mean change (SE) 1) -14.1 (0.6) 2) -7.7 (1.2), *p<0.*001 vs 1 3) -11.7 (1.1), *NS* vs 1 4) -13.7 (1.2), *NS* vs 1 5) -12.0 (1.1), *NS* vs 1	Response (MADRS), % 1) 43.3 2) 25.4, *p=0.017 vs 1* 3) 33.9, *NS* vs 1 4) 50.0, *NS* vs 1 5) 36.4, *NS* vs 1 Remission: MADRS≤8 at 2 consecutive visits, % 1) 29.9 2) 13.6, *p=0.013* vs 1 3) 17.9, *NS* vs 1 4) 22.4, *NS* vs 1 5) 20.0, *NS* vs 1	OFC-treated patients reported higher rates of somnolence and peripheral edema compared to VNL and FLX (all p<0.05). Rates of weight gain and increased appetite were higher in OFC-than VNL-treated patients (both p<0.05). D/C due to weight gain occurred at a higher incidence in OFC-treated patients compared to other groups.
rTMS						
O'Reardon 2007[75] Multicenter, North America & Australia 4-6 week double-blind phase. At week 4, patients with <25% improvement	1) Sham (n=146) 2) Unilateral rTMS (n=155) Concomitant AD not allowed	Failed 1-4 ADs in current or most recent episode or history of intolerance to at least 4 ADs	– Mean age: 48.3 years – Mean MADRS: 33.3 – Mean $HAMD_{17}$: 22.7 – Mean duration of current episode: 1.1 years – Mean AD failures in current episode: 1.6	At week 4: MADRS, est. mean change§ 1) -4.1 2) -5.8, *NS* vs 1 $HAMD_{17}$, est. mean change§ 1) -3.5 2) -5.2, *p=0.006* vs 1	At week 4: Response (MADRS), % 1) 11.0 2) 18.1, *p<0.05* vs 1 Response ($HAMD_{17}$), % 1) 11.6 2) 20.6, *p<0.05* vs 1 Remission: MADRS≤9, % 1) 6.2	More patients treated with rTMS reported scalp discomfort and pain compared to sham. Nine and seven SAEs in the rTMS and sham groups were reported, respectively. Most SAEs were disease-

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
could crossover and receive open-label treatment *Poor quality*					2) 7.1, *NS* vs 1 Remission: $HAMD_{17}$≤7, % 1) 6.2 2) 7.1, *NS* vs 1	related.
George 2010[70] Multicenter, US 3-6 week double-blind phase. At week 3, improvers (>30% improvement) who have yet to remit received up to 3 weeks of additional tx *Fair quality*	1) Sham (n=105) 2) Unilateral rTMS (n=94) Concomitant AD not allowed	Insufficient clinical benefit to 1-4 ADs, or intolerant to ≥3 ADs (not specified in in current episode or history)	– Mean age: 47.1 years – Mean MADRS: 29.6 – Mean $HAMD_{24}$: 26.4 – Mean duration of current episode: 1.5 years – Mean AD failures in current episode: NR* *No. of AD failures not reported, but mean no. of failed research-quality antidepressant trials assessed by Antidepressant Treatment History Form [111] was 1.5	At week 6: MADRS, est. mean change§ 1) -2.1 2) -4.9, *p=0.01 vs 1* $HAMD_{24}$, est. mean change§ 1) -3.1 2) -4.7, *NS vs 1*	Response ($HAMD_{24}$), % At week 6: 1) 5 2) 15, *p=0.009 vs 1* Remission: $HAMD_{24}$≤3 or ≤10 at 2 consecutive visits, % At week 3: 1) 2 2) 6.5, *p NR* At endpoint (week 3-6): 1) 5 2) 14, *p=0.02 vs 1*	More patients treated with rTMS reported headache and discomfort at administration site compared to sham. Five patients, all receiving rTMS, discontinued due to AEs; four of the five patients reported pain or headache as the reason for d/c.
Pallanti 2010[79] Single Center, Italy 3-week double-blind phase	1) Sham + AD (n=20) 2) Unilateral rTMS + AD (n=20) 3) Bilateral rTMS + AD (n=20) All continued AD (SSRI, SNRI, TCA, bupropion)	Failed ≥2 classes of ADs in history according to Thase and Rush criteria[112]	– Mean age: 48.9 years – Mean $HAMD_{17}$: 28.6 – Mean duration of current episode: 0.8 years – Mean AD failures in history*: 5.9	$HAMD_{17}$, est. mean change§ 1) -2.2 2) -6.9 3) -10.7 *p-values NR*	Response ($HAMD_{17}$), % 1) 10 2) 35 3) 20 *p<0.05 among all groups. Pairwise p-values NR.* Remission: $HAMD_{17}$≤8 %	During the first week, more patients treated with rTMS reported scalp pain and headache compared to sham. However, by week 3, the incidence of

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
Fair quality			*Failures in current episode NR		1) 5 2) 30 3) 10 *p NS among all groups. Pairwise p-values NR.*	headache and scalp pain were similar in the rTMS and sham groups.
Bakim 2012[77] Single Center, Turkey 6-week double-blind phase *Poor quality*	1) Sham + AD (n=12) 2) Unilateral 80% rTMS + AD (n=12) 3) Unilateral 110% rTMS + AD (n=11) All continued AD (SSRI or SNRI)	Lack of clinically significant decrease in depressive symptoms to ≥2 different classes of ADs in current episode	– Mean age: 42.1 years – Mean MADRS: 27.9 – Mean $HAMD_{17}$: 24.3 – Duration of last episode*: 1.4 years – Mean AD failures in current episode: 3.5 *Duration of current episode NR	MADRS, percent change 1) 25.1% 2) 58.6%, *p=0.01 vs 1* 3) 50.7%, *p=0.05 vs 1* $HAMD_{17}$, percent change 1) 25.8% 2) 58.3%, *p=0.01 vs 1* 3) 52.0% *p=0.04 vs 1*	Response (MADRS), % 1) 16.7 2) 75.0, *p=0.01* vs 1 3) 72.7, *p=0.01* vs 1 Response ($HAMD_{17}$), % 1) 16.7 2) 83.3, *p<0.01* vs 1 3) 72.7, *p=0.01* vs 1 Remission: $HAMD_{17}$≤7, % 1) 8.3 2) 25.0, *NS* vs 1 3) 54.5, *p=0.03* vs 1	No SAEs were reported during the study. Four participants (17.4%) receiving active rTMS and one participant (8.3%) treated with sham reported mild headaches during the study. Two participants (8.7%) treated with active sham reported mild discomfort at the administration site.
Blumberger 2012[74] Single Center, Canada 3-6 week double-blind phase. At week 3, non-remitters received up to 3	1) Sham ± AD (n=22) 2) Unilateral rTMS ± AD (n=24) 3) Bilateral rTMS ± AD (n=28) 60% of patients continued their AD during trial	Failed to achieve clinical response or did not tolerate ≥2 classes of ADs in current episode according to Thase and Rush criteria[112]	– Mean age: 51.5 years – Mean $HAMD_{17}$: 25.4 – Duration of current episode: NR – Mean AD failures in current episode: NR	$HAMD_{17}$, est. mean change§ At week 3: 1) -7.4 2) -6.4 3) -9.8 *p-values NR* $HAMD_{17}$, percent change At week 6: 1) 24.9% 2) 23.0%, *NS vs 1*	Response ($HAMD_{17}$), % At endpoint (week 3-6): 1) 10 2) 4.5, *NS* vs 1 3) 38.5, *p=0.022* vs 1 Remission $HAMD_{17}$≤10, % At week 3: 1) 0 2) 4.5 3) 15.4	Three patients discontinued after experiencing SAEs judged unrelated to study treatment (myocardial infarction in bilateral group and suicidality requiring hospitalization in unilateral and sham

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
weeks of additional tx *Poor quality*				3) 44.0%7, *p=0.032 vs 1*	*p-values NR* At endpoint (week 3-6): 1) 5 2) 4.5, *NS* vs 1 3) 34.6, *p=0.028* vs 1	groups). One patient in the unilateral group also withdrew due to insomnia.
Fitzgerald 2012[73] Single Center, Australia 3-week double-blind phase *Poor quality*	1) Sham ± AD (n=20) 2) Unilateral rTMS ± AD (n=24) 3) Bilateral rTMS ± AD (n=22) 61% of patients continued their AD during trial	Failed to respond to ≥2 classes of ADs in history according to Thase and Rush criteria[112]	– Mean age: 42.9 years – Mean MADRS: 32.5 – Mean $HAMD_{17}$: 23.7 – Duration of current episode: NR – Mean AD failures in history*: 5.2 *Failures in current episode NR	MADRS, est. mean change§ 1) -2.0 2) -4.5 3) -2.5 *p NS among all groups. Pairwise p-values NR.* $HAMD_{17}$, est. mean change§ 1) -0.2 2) -4.1, *p=0.02* vs 1 3) -2.1, *NS* vs 1	Response ($HAMD_{17}$), % 1) 0 2) 0 3) 5 *p-values NR*	There were no SAEs reported during the trial.
Chen 2013[78] Single Center, China 4-week double-blind phase *Poor quality*	1) Sham ± AD (n=10) 2) Unilateral rTMS ± AD (n=10) All continued AD (NR)	Failed to respond to 2 ADs (not specified in in current episode or history)	– Mean age: 45.7 years – Mean $HAMD_{17}$: 24.2 – Duration of current episode: NR – No. AD failures in current episode: NR	$HAMD_{17}$, est. mean change§ 1) -12.6 2) -13.9 *p-value NR*	Response ($HAMD_{17}$), % 1) 80 2) 70, *NS* vs 1	One patient in the sham group discontinued due to unspecified somatic issues.
Brunelin 2014[72] Multicenter, France & Monaco	1) Unilateral rTMS ± Venlafaxine [Combo] (n=55) 2) Sham + Venlafaxine [VNL] (n=55)	Persisting depressive symptoms ($HAMD_{17}$)>20 despite receiving	– Mean age: 54.5 years – Mean $HAMD_{17}$: 25.9 – Mean MADRS: 33.0 – Mean duration of	MADRS, est. mean change§ At week 2: 1) -6.9 2) -7.8 3) -7.2	Response ($HAMD_{17}$), % At endpoint (2-6 weeks): 1) 54 2) 60 3) 59	Twelve SAEs were reported: seven in VNL group, three in the combo group, and two in the rTMS

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
2-6 week double-blind phase. Patients continued to receive treatment until remission. *Good quality*	3) Unilateral rTMS + Placebo [rTMS] (n=60) VNL and combination groups initiated new AD. rTMS arm did not receive concomitant AD.	treatment with AD	current episode: 1.4 years – Mean AD failures in current episode: 2.5	At week 6: 1) -13.4 2) -14.7 3) -14.9 $HAMD_{17}$, est. mean change§ At week 2: 1) -6.0 2) -6.0 3) -6.2 At week 6: 1) -10.7 2) -11.5 3) -11.8 *p-values NR*	*p NS among all groups. Pairwise p-values NR.* Remission: $HAMD_{17}$≤7, % At week 2: 1) 2.0 2) 3.9 3) 7.4 *p-values NR* At endpoint (2-6 weeks): 1) 28.0 2) 43.1 3) 40.7 *p NS among all groups. Pairwise p-values NR.*	group. The most common AE, occurring in five patients, was exacerbation of depressive symptoms leading to hospitalization.
Blumberger 2016[80] Single Center, Canada 3-6 week double-blind phase. Patients continued to receive treatment until remission. *Poor quality*	1) Sham ± AD (n=41) 2) Unilateral rTMS ± AD (n=40) 3) Bilateral rTMS ± AD (n=40) 95% of patients continued their AD during trial	Failed to achieve clinical response or did not tolerate ≥2 classes of ADs in history according to Thase and Rush criteria[112]	– Mean age: 47.0 years – Mean $HAMD_{17}$: 25.2 – Mean duration of current episode: 3.6 years – No. of AD failures in current episode: NR* *No. of AD failures NR, but the mean Antidepressant History Treatment Form[111] score was 7.4.	$HAMD_{17}$, mean change 1) -5.0 2) -6.4 3) -6.8 *p NS among all groups*	Response ($HAMD_{17}$), % At endpoint (3-6 weeks): 1) 4.9 2) 15, *NS vs 1* 3) 22.5, *p=0.026* vs 1 Remission: $HAMD_{17}$≤7, % A week 3: 1) 2.4 2) 0 3) 7.5 *p-values NR* At endpoint (3-6 weeks):	The most commonly reported AE was headache which occurred at similar rates in all three groups. More patients treated with rTMS reported pain compared to those treated with sham.

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
					1) 2.4 2) 7.5, *NS* vs 1 3) 20, *p=0.014* vs 1	
Carpenter 2017[71] Multicenter, US 4-6 week double-blind phase. Patients completed treatment protocol within 4-6 weeks. *Fair quality*	1) Sham ± AD (n=45) 2) Bilateral rTMS ± AD (n=47) 64% of patients continued their AD during trial	Failed to receive benefit from 1-3 ADs in current episode or did not tolerate ≥1 AD in current or past episode according to MGH-ATRQ	– Mean age: 46.6 years – Mean $HAMD_{24}$: 31.2 – Mean duration of current episode: 1.4 years – Mean AD failures in current episode: 1.3	$HAMD_{24}$ mean change (SD) 1) -10.4 (8.7) 2) -15.1 (9.6), *p=0.03* vs 1	Response ($HAMD_{24}$) 1) 32.4 2) 55.3, *NS* vs 1 Remission: $HAMD_{24}$≤10,% 1) 18.9 2) 26.3, *NS* vs 1	Headache and pain at the administration site were more frequently reported by patients treated with rTMS compared to those receiving sham. No SAEs were observed in the rTMS group. Two SAEs (suicide attempt and hypotensive event) were reported in the sham group.
Theleritis 2017[76] Single Center, Greece 3-week double-blind phase with 2-week follow-up *Poor quality*	1) Sham ± AD once per day 2) rTMS ± AD once per day 3) Sham ± AD twice per day 4) rTMS ± AD twice per day For groups 1 & 3, n=44. For groups 2 & 4 n=54. 56% of patients continued their AD during trial	Failed to response to ≥2 classes of ADs in history according to Thase and Rush criteria[112]	– Mean age: 38.9 years – Mean $HAMD_{17}$: 30.0 – Duration of current episode: NR – No. of AD failures in current episode: NR* * No. of AD failures NR, but about 25% of patients had Thase and Rush stage ≥4.[112]	$HAMD_{17}$, est. mean change§ At week 3: 1) -4.0 2) -15.0 3) -3.3 4) -16.6 At week 5:* 1) -3.5 2) -15.7 3) -2.9 4) -17.4 *p-values NR* **Assessment was extended*	At week 5: Response ($HAMD_{17}$), % 1) & 3): 2.5 2) & 4): 59.2, *p<0.001* vs 1 & 3 Remission: $HAMD_{17}$≤7, % 1) & 3): 0 2) & 4): 24.5, *p=0.001* vs 1 & 3 *Response and remission rates were not reported at week 3 (end of double-*	A similar proportion of patients receiving sham and rTMS reported discomfort at administration site and exacerbation of preexisting headache during the trial. One patient receiving rTMS and one patient receiving sham discontinued due to exacerbation of preexisting

	Study Arms (n), Concomitant AD	Definition of Treatment-Resistance	Population Characteristics	Change From Baseline in Depressive Scale*	Response† and Remission‡	Harms
				2 weeks beyond end of double-blind phase because of the possibility of a late onset of effect	*blind phase)*	headache.
ECT vs. TMS						
Rosa 2006[82] Single Center, Brazil 4-week double-blind phase *Poor quality*	1) ECT (n=20) 2) Unilateral rTMS (n=22) Concomitant AD not allowed	Lack of response to ≥2 classes of AD, with augmentation for ≥1 trial (not specified if in current episode or history)	– Mean age: 43.1 years – Mean $HAMD_{17}$: 31.0 – Mean duration of current episode: 0.9 years – No. of AD failures in current episode: NR	$HAMD_{17}$, est. mean change§ 1) -13.6 2) -12.7 *p-values NR*	Response ($HAMD_{17}$), % 1) 30 2) 45, *NS* vs 1 Remission: $HAMD_{17}$≤7, % 1) 15 2) 9, *NS* vs 1	NR

AD: antidepressant, AE: adverse event, D/C: discontinuation, ECT: Electroconvulsive therapy, $HAMD_{17}$: Hamilton Depression Rating Scale 17-item, $HAMD_{24}$: Hamilton Depression Rating Scale 24-item, MADRS: Montgomery–Åsberg Depression Rating Scale, MGH-ATRQ: Massachusetts General Hospital-Antidepressant Treatment Response Questionnaire; n: number of patients at randomization, NR: not reported, NS: not significant, PBO: Placebo; rTMS: repetitive transcranial magnetic stimulation, SAE: serious adverse event, SD: standard deviation, SE: standard error, SNRI: Serotonin Noradrenaline Reuptake Inhibitor, SSRI: Selective Serotonin Reuptake Inhibitor, TEAE: treatment-emergent adverse event, TCA: tricyclic antidepressant, tx: treatment

*Change from baseline to end of double-blind period, unless otherwise stated. †Percentage of patients with ≥50% reduction in depressive rating scale from baseline to end of double-blind period, unless otherwise stated, ‡Percentage of patients with score at or below threshold at end of double-blind period unless otherwise stated, §Estimated mean change was calculated when depressive rating score was given at baseline and endpoint

Appendix E. Comparative Value Supplemental Information

Table E1. Impact Inventory

Sector	Type of Impact (Add Additional Domains, as Relevant)	Included in This Analysis From… Perspective?		Notes on Sources (If Quantified), Likely Magnitude & Impact (If Not)
		Health Care Sector	Societal	
Formal Health Care Sector				
Health Outcomes	Longevity effects	X	X	
	Health-related quality of life effects	X	X	
	Adverse events	X	X	
Medical Costs	Paid by third-party payers	X	X	
	Paid by patients out-of-pocket	☐	☐	
	Future related medical costs	☐	☐	
	Future unrelated medical costs	☐	☐	
Informal Health Care Sector				
Health-Related Costs	Patient time costs	NA	☐	
	Unpaid caregiver-time costs	NA	☐	
	Transportation costs	NA	☐	
Non-Health Care Sectors				
Productivity	Labor market earnings lost	NA	X	
	Cost of unpaid lost productivity due to illness	NA	X	
	Cost of uncompensated household production	NA	☐	
Consumption	Future consumption unrelated to health	NA	☐	
Social Services	Cost of social services as part of intervention	NA	☐	
Legal/Criminal Justice	Number of crimes related to intervention	NA	☐	
	Cost of crimes related to intervention	NA	☐	
Education	Impact of intervention on educational achievement of population	NA	☐	
Housing	Cost of home improvements, remediation	NA	☐	
Environment	Production of toxic waste pollution by intervention	NA	☐	
Other	Other impacts (if relevant)	NA	☐	

NA: not applicable

Adapted from Sanders et al., 2016[113]

Table E2. Selected Listing of Model Transition Probabilities and Conversions from Sources

Model Probability	Description	Calculation or Calculation Components	Value	Source
Transition From Initial Esketamine Treatment Severe Depression to Initial Esketamine No Depression	Those patients in whom esketamine was effective. Weighted average of placebo remission proportion from TRANSFORM-1 and 2 multiplied by RR for esketamine remission from meta-analysis of TRANSFORM-1 and 2	Placebo weighted proportion achieving remission = 28.8%; RR = 1.37 (0.99-1.91)	0.3947	TRANSFORM-1 and 2 meta-analysis
Transition From Initial Esketamine Treatment Severe Depression to Initial Esketamine Mild Depression	Those patients in whom esketamine was partly effective. Calculated as the weighted placebo response rate from the TRANSFORM-1 and 2 trials multiplied by the calculated response ratio from the TRANSFORM-1 and 2 meta-analysis, minus the weighted proportion who achieved remission on esketamine from the TRANSFORM-1 and 2 trials.	Weighted placebo response = 0.453; Relative ratio = 1.3 (1.08-1.56); weighted esketamine remission rate = 0.3947. Value=(0.453*1.3)-0.3947	0.1942	TRANSFORM-1 and 2 meta-analysis
Transition From Initial Esketamine Treatment Severe Depression to Alternative Treatment 1 No Depression	Those patients in whom esketamine was not effective and the alternative treatment is effective. Calculated as the proportion of patients not receiving either full or partial effect (p1 and p8) or dying (p39) multiplied by the probability of a step4 therapy remission from the STAR*D study table 4.	STAR*D remission probability for step 4 is 0.13; the equation is (1-(p1+p8+p38))*0.13 where p1 and p8 are the transitions to effective and partly effective treatment and p38 is the age dependent mortality rate.	0.0536 for the first cycle	Calculated; TRANSFORM-1,2 MA, STAR*D table 4
Transition From Initial Esketamine Treatment Severe Depression to Alternative Treatment 1 Severe Depression	Those patients in whom esketamine was not effective and the alternative treatment is not effective. Calculated as the proportion of patients not receiving either full or partial effect (p1 and p8) or dying (p39) multiplied by the probability of a step4 therapy failure to achieve remission from the STAR*D study table 4.	STAR*D remission probability for step 4 is 0.13; equation is (1-(p1+p8+p38))*(1-0.13) where p1 and p8 are the transitions to effective and partly effective treatment and p38 is the age dependent mortality rate.	0.3587 for the first cycle	Calculated; TRANSFORM-1,2 MA, STAR*D table 4
Transition From Initial Esketamine No Depression to Initial Esketamine Discontinued No Depression	Those who received effect from initial treatment and had their initial treatment discontinued because of long term effectiveness	Expert opinion; 5% per year, using appropriate version of calculation p=1-e^-kt to convert from one-year probability to 3-month probability.	0.0127	Expert opinion
Transition From Initial Esketamine	Those who had initial effect on esketamine, but subsequently lost	Relapse on esk remitters from SUSTAIN poster	0.1148	SUSTAIN-1 (poster figure

Model Probability	Description	Calculation or Calculation Components	Value	Source
No Depression to Alternative Treatment 1 No Depression	effect and did not receive effect from the alternative treatment.	figure 2 (1-0.868) multiplied by no remission on step5 from STAR*D table 4 (1-0.13)		2a), StarD table 4
Transition From Initial Esketamine No Depression to Alternative Treatment 1 Severe Depression	Those who had initial effect on esketamine, but subsequently lost effect and received effect from the alternative treatment.	Relapse on esk remitters from SUSTAIN poster figure 2 (1-0.868) multiplied by remission on step5 from STAR*D table 4 (0.13)	0.0172	SUSTAIN-1 (poster figure 2a), StarD table 4
Transition From Initial Esketamine Mild Depression to Initial Esketamine No Depression	Those patients in whom esketamine was initially partly effective and then became effective. Reported from SUSTAIN-1 Poisson regression in open input.	Value = 0.199	0.199	SUSTAIN-1 (Poisson regression, reported from open input document page 10)
Transition From Initial Esketamine Mild Depression to Alternative Treatment 1 No Depression	Those patients who lose partial effect with esketamine, but subsequently lost effect and received effect from the alternative treatment. Calculated as relapse on esketamine in those with response from SUSTAIN-1 poster multiplied by those in with remission on step5 from STAR*D	Relapse on esketamine in responders (1-0.79)*remission on step5 from StarD (0.13)	0.0273	SUSTAIN-1 (poster figure 2b), StarD table 4
Transition From Initial Esketamine Mild Depression to Alternative Treatment 1 Severe Depression	Those patients in whom esketamine was partially effective, but subsequently lost effect and did not receive effect from the alternative treatment. Calculated as relapse on esketamine in those with response from SUSTAIN-1 poster multiplied by those in with remission on step5 from STAR*D.	Relapse on esketamine in responders (1-0.79)*no remission on step5 (1-0.13) from StarD	0.1827	"SUSTAIN-1 (poster figure 2b), StarD table 4

Table E3. Undiscounted Base-Case Results

Treatment Pathways	Drug Cost	Total Cost	QALYs	LYs	Depression-Free Day
Esketamine	$43,500	$718,200	20.73	34.01	240
No Additional Treatment	$0	$678,000	20.52	33.98	119
Difference	$43,500	$38,400	0.21	0.03	121

QALY: quality-adjusted life year, LY: life year

Table E4. Esketamine Versus No Additional Treatment: Cost-Effectiveness Acceptability Curve

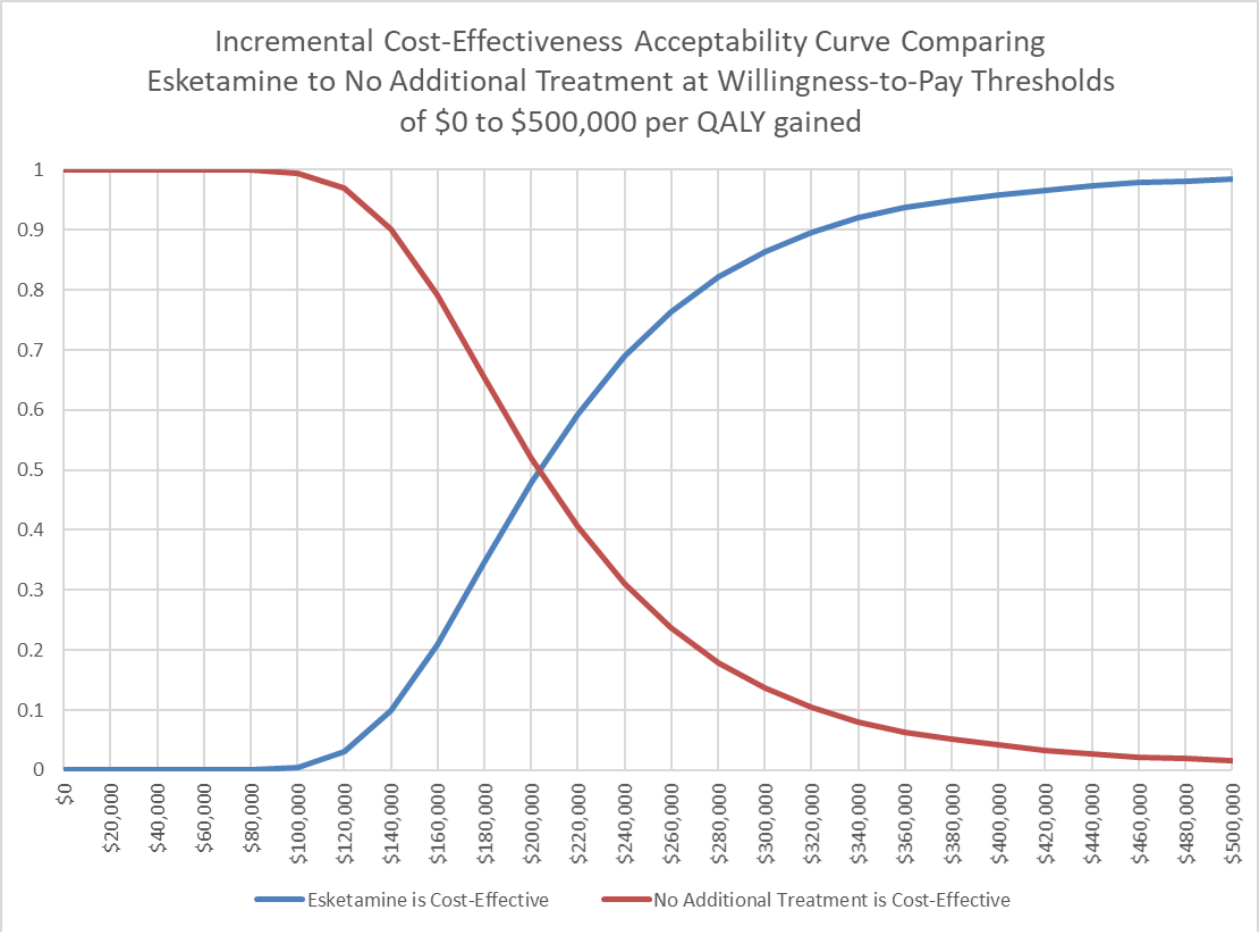


Table E5. Scenario Analysis Results: Modified Societal Perspective

Treatment Pathways	Drug Cost	Total Cost	QALYs	LYs	Depression-Free Day
Esketamine	$42,600	$422,700	12.66	20.66	235
No Additional Treatment	$0	$386,300	12.47	20.64	117
Difference	$42,600	$36,300	0.19	0.01	117

QALY: quality-adjusted life year, LY: life year