Vesicular Monoamine Transporter 2 Inhibitors for Tardive Dyskinesia: Effectiveness and Value

Draft Evidence Report

October 2, 2017

Prepared for

ICER Staff/Consultants	University of Illinois at Chicago College of Pharmacy Modeling Group*
Steven J. Atlas, MD, MPH Director, Primary Care Research and Quality Improvement Network Massachusetts General Hospital	**Surrey M Walton, PhD** Associate Professor of Pharmacy Department of Pharmacy Systems, Outcomes, and Policy University of Illinois at Chicago College of Pharmacy
Foluso Agboola, MBBS, MPH Research Scientist Institute for Clinical and Economic Review	**Daniel R. Touchette, PharmD, MA, FCCP** Associate Professor of Pharmacy and Takeda Professor of Medication Adherence Department of Pharmacy Systems, Outcomes, and Policy University of Illinois at Chicago College of Pharmacy
Ifeoma Otuonye, MPH Research Associate Institute for Clinical and Economic Review	**Katherine Harrigan, PharmD** Research Fellow Department of Pharmacy Systems, Outcomes, and Policy University of Illinois at Chicago College of Pharmacy
Varun Kumar, MBBS, MPH, MSc Health Economist Institute for Clinical and Economic Review	
Celia Segel, MPP Program Manager Institute for Clinical and Economic Review	
Rick Chapman, PhD, MS Director of Health Economics Institute for Clinical and Economic Review	
Daniel A. Ollendorf, PhD Chief Scientific Officer Institute for Clinical and Economic Review	*The role of the University of Illinois at Chicago (UIC) College of Pharmacy Modeling Group is limited to the development of the cost-effectiveness model, and the resulting ICER reports do not necessarily represent the views of the UIC.
Steven D. Pearson, MD, MSc President Institute for Clinical and Economic Review	

DATE OF PUBLICATION: October 2, 2017

We would also like to thank Aqsa Mugal, Molly Morgan, and Erin Lawler for their contributions to this report.

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. ICER receives funding from government grants, non-profit foundations, health plans, provider groups, and health industry manufacturers. For a complete list of funders, visit http://www.icer-review.org/about/support/. 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

About New England CEPAC

The New England Comparative Effectiveness Public Advisory Council (New England 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. New England CEPAC seeks to help patients, clinicians, insurers, and policymakers interpret and use evidence to improve the quality and value of health care.

The New England CEPAC is an independent committee of medical evidence experts from across New England, 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 New England CEPAC is available at http://icer-review.org/programs/new-england-cepac/.

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/wp-content/uploads/2017/05/NECEPAC_TD_Stakeholder_List_05082017.pdf

Expert Reviewers

Robert Rosenheck, MD
Department of Psychiatry
Yale Medical School
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers

Cindy Specht
Executive Vice President
Depression and Bipolar Support Alliance (DBSA)
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers.

Daniel Tarsy MD
Beth Israel Deaconess Medical Center
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers

Table of Contents

List of Acronyms Used in this Report

AAN	American Academy of Neurology
AHRQ	Agency for Healthcare Research and Quality
AIMS	Abnormal Involuntary Movement Scale
AE	Adverse event
BID	Twice a day
CGIC	Clinical Global Impression of Change
DBPC	Double blind placebo controlled
DRBAs	Dopamine receptor blocking agents
DTBZ	Deutetrabenazine
ESRS	Extrapyramidal Symptoms Rating Scale
FDA	Food and Drug Administration
ITT	Intent-to-treat
LTE	Long term extension
MAOIs	Monoamine oxidase inhibitors
mCDQ	Modified Craniocervical Dystonia Questionnaire
mITT	Modified intent-to-treat population
NNH	Number needed to harm
NNT	Number needed to treat
OLE	Open-label extensions
PBO	Placebo
PGIC	Patient Global Impression of Change
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY	Quality-Adjusted Life Year
RCT	Randomized Control Trial
TBZ	Tetrabenazine
TD	Tardive dyskinesia
TEAEs	Treatment emergent adverse events
TID	Three times a day
USPSTF	U.S. Preventive Services Task Force
VBZ	Valbenazine
VMAT2	Vesicular monoamine transporter-2

1. Background

1.1 Introduction

Background

Tardive dyskinesia (TD) is a repetitive, involuntary movement disorder with a delayed onset caused by prolonged use of medications that block the dopamine receptor.[1] The most common medications causing TD are antipsychotic drugs used to treat individuals with schizophrenia and schizoaffective disorder, mood disorders such as bipolar disease and major depression, and dementia, personality disorders, post-traumatic stress disorder, and insomnia.[2,3] Less commonly-used medications associated with TD include metoclopramide for gastrointestinal disorders and certain antidepressants with dopamine receptor blocking properties such as amoxapine.[4,5]

Overall, 20-30% of individuals taking antipsychotic drugs develop TD, with increased prevalence associated with longer duration of use, older age, female gender, and the specific antipsychotic drug used.[6,7] Newer antipsychotic drugs, referred to as second-generation or atypical agents, are somewhat less likely to cause TD than older, first-generation antipsychotic drugs.[8] The incidence of developing TD is reported to be around 5% per year with first-generation and 3% per year with second-generation antipsychotics.[9,10] However, the increased use of newer antipsychotic drugs mean that they represent most new cases of TD.

The movements associated with TD can be localized or widespread and can result in physical and psychological impairment.[11] TD commonly involves the mouth and face region with lip smacking or pursing, chewing, facial grimacing, and tongue movements. TD can also affect the head and neck with abnormal asymmetrical postures, and the limbs and trunk with foot tapping, finger movements, and shoulder shrugging. Patients may not be aware of these movements, especially when involving the face. For most, symptoms are mild to moderate. When severe, TD can lead to difficulty swallowing, trouble walking, and impaired speech. Thus, TD can be extremely debilitating and result in social isolation, problems in the workplace, and decreased compliance with the drugs given to treat the underlying condition.[12]

Many treatments have been tried for TD, but there has been limited evidence generated to date that can guide patients and clinicians.[2,13,14] Although reducing or discontinuing the offending agent would seem logical, this is often not possible because TD symptoms can paradoxically increase on withdrawal and the underlying condition may worsen. Until recently there were no FDA-approved therapies for TD. However, the approval by the FDA of novel vesicular monoamine transporter-2 (VMAT2) inhibitors, valbenazine and deutetrabenazine, has led to new interest in TD and its treatment.[15]

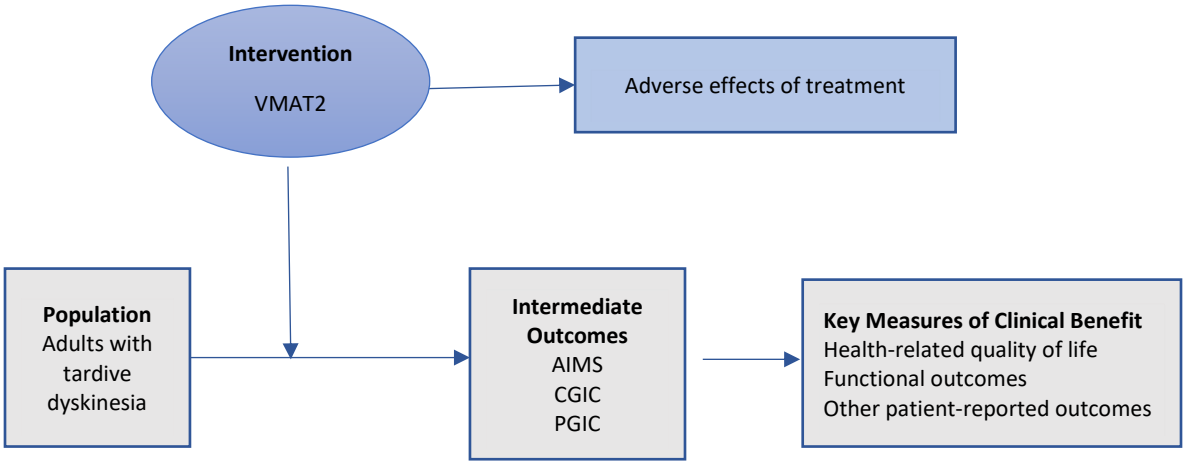
Scope of the Assessment

This review evaluated the comparative clinical effectiveness of the VMAT2 inhibitors valbenazine, deutetrabenazine, and tetrabenazine for the treatment of adults with TD. 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; however, when assessing adverse events and harms, we also looked for randomized trials of the VMAT2 inhibitors for conditions other than TD. We did not restrict studies according to study duration or study setting; however, we limited our review to those that measured the outcomes of interest and included at least 10 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 meets ICER standards (for more information, see http://icer-review.org/methodology/icers-methods/icer-value-assessment-framework/grey-literature-policy/). We sought out head-to-head studies of these interventions in order to consider combined use of direct and indirect evidence in network meta-analyses of selected outcomes.

Analytic Framework

The analytic framework for this assessment is depicted in Figure 1.1.

Figure 1.1. Analytic Framework: Management of Tardive Dyskinesia



VMAT2: vesicular monoamine transporter 2; AIMS= Abnormal Involuntary Movement Scale; CGIC= Clinical Global Impression of Change; PGIC= Patient Global Impression of Change

Populations

The primary population of focus for the review was adults ages 18 and older with symptoms of TD for at least three months and history of use of dopamine receptor blocking agents (DRBAs).[16] In addition to children and adolescents, we also excluded studies of adults with movement disorders related to other conditions (e.g., Huntington's disease) or whose disorder was not thought to be medication-induced. However, populations with conditions other than TD that use the interventions of interest were also assessed when examining adverse events and other potential harms.

We also sought evidence on key subpopulations of interest, including: (a) patients with incident or new onset TD; (b) patients with localized versus generalized TD symptoms; (c) patients with schizophrenia/schizoaffective disorders versus other underlying conditions that may be associated with TD. Other subgroups of interest included age, gender, and severity of symptoms as assessed by both clinicians and patients (i.e., mild, moderate, or severe).

Interventions

We considered all VMAT2 inhibitors including those with FDA indications for TD as well as one drug that is used off-label for TD. Interventions of interest are listed below.

- Valbenazine (Ingrezza™, Neurocrine Biosciences, Inc)
- Deutetrabenazine (Austedo®, Teva)
- Tetrabenazine (Xenazine®, Lundbeck and generics [off-label use])

Comparators

We examined studies comparing VMAT2 inhibitors to placebo or other types of active medications that are used off-label to control TD symptoms. Wherever possible, we evaluated head-to-head trials of the interventions. If suitable data was available, the review sought to include head-to-head comparisons through methods such as network meta-analysis.

Outcomes

This review examined key clinical outcomes associated with TD. However, when assessing adverse events and harms, we also looked for randomized trials of the interventions of interest for conditions other than TD. We engaged with patient groups and clinical experts to ascertain which outcomes are of greatest importance to patients, and sought patient-reported outcomes or other evidence sources to enrich the available data. Discussion with patients, patient groups, and clinicians indicated that clinical trials often lack robust information on patient-reported outcomes and burdens associated with TD.

Outcomes of interest included:

- Symptom improvement (Abnormal Involuntary Movement Scale [AIMS], Clinical Global Impression of Change [CGIC])
- Patient reported outcome (Patient Global Impression of Change [PGIC])
- Health-related quality of life
- Treatment-related adverse events (e.g., somnolence, suicide, worsening of underlying mental health illness)
- Discontinuation due to adverse events
- Costs and cost-effectiveness

We also sought evidence on additional patient-reported outcomes as available. Importantly, long-term use of antipsychotic drugs is also associated with the development of other extrapyramidal symptoms and movement disorders, but the focus of this assessment was on TD symptoms only.

Evidence tables were developed for each selected study and results were summarized in a qualitative fashion. If available data permitted, we sought to perform meta-analysis to quantitatively summarize outcomes for the therapies of interest, and network meta-analysis to combine direct and indirect evidence of effectiveness.

Timing

Evidence on intervention effectiveness was derived from studies of any duration if they meet the study design criteria set forth above and measure the outcomes of interest.

Settings

All relevant settings were considered, including outpatient/clinic, office, and home settings.

2. The Topic in Context

TD was first described in the 1950s with the introduction of first-generation antipsychotic drugs (e.g., chlorpromazine).[3] Recognition that TD was a drug-induced condition caused by long-term use of antipsychotic drugs led to the introduction of clozapine and other newer antipsychotic drugs (also called atypical or second-generation antipsychotic drugs) starting in the 1970s, with the goal of decreasing TD and other abnormal movements attributed to these medications. Antipsychotic drugs are believed to work in the brain as dopamine receptor blocking agents (DRBAs), and the newer antipsychotic drugs have properties that were intended to decrease the likelihood of TD and other abnormal movements while maintaining the drugs' beneficial effects. Though antipsychotic drugs are the most common DRBAs used in medical practice, other DRBAs such as metoclopramide, which is used to treat certain gastrointestinal problems that cause nausea and vomiting, can also cause TD.

2.1 How Common is Tardive Dyskinesia?

TD is common in patients treated with DRBAs for long periods of time. The prevalence rates of TD have been estimated to range from 20-50%, and is thought to be higher for first-generation (30%) than for second-generation (13-20%) agents.[6,17] The incidence of new TD is reported to be around 5% per year with first-generation antipsychotic drugs and 3% per year with second-generation antipsychotic drugs.[9,10,18] Higher rates are seen with longer duration of use and in female and older patients.[7] Though newer antipsychotic drugs may have a somewhat lower risk of TD, rates are not as low as initially hoped, and the use of these agents have increased with an expanding range of off-label indications.[19,20]

While antipsychotic drugs were initially developed for use in patients with schizophrenia and schizoaffective disorder, they are also used in serious mood disorders such as bipolar disease and major depression. Off-label use of antipsychotics include treatment of obsessive-compulsive disorder, personality disorders, post-traumatic stress disorder, insomnia, and agitation in dementia. Symptoms of TD can vary from mild to severe and can adversely affect all aspects of patient's lives including physical and psychological impairment, social isolation, and work-related disability. It is estimated that there are 6 million individuals in the United States with these diagnoses who have received antipsychotic drugs.[21] Among these persons, there may be 500,000 individuals with TD if the prevalence is 25%, for example. For most, TD symptoms are irreversible, even if the DRBA is stopped.[22] Younger individuals and those on DRBAs for shorter periods of time may have decreased TD symptoms with cessation, but complete resolution is uncommon. If new treatments are targeted to those with moderate or severe TD and 60% have such symptoms, then one may estimate that 300,000 patients would be eligible for new treatments.

2.2 How is Tardive Dyskinesia Diagnosed?

The diagnosis of TD is based upon presence of abnormal, involuntary movements, a history of taking a DRBA for at least three months, and the absence of another condition that may produce such movements. Diagnosis is therefore clinical in nature only, meaning there are no specific tests that can be done to confirm or "prove" the diagnosis. The severity of TD symptoms and the impact of treatment can be assessed using survey questionnaires. The most common is the Abnormal Involuntary Movement Scale (AIMS) that has been used in clinical and research settings.[23] The AIMS score is based on the measurement of seven items (four involving facial/oral movements, two involving the extremities, and one the trunk), with severity rated from 0 (none) to 4 (severe). The total score ranges from 0-28 with higher AIMS scores reflecting increased severity of TD symptoms. (see Appendix Table F1). While the AIMS assesses symptom severity, it is unclear how well it reflects the overall impact of TD on a patient's quality of life.

The abnormal, involuntary movements seen in TD are slow, rhythmical, and stereotyped, and can be generalized or limited to specific parts of the body. Most common is involvement of the mouth and face region with abnormal movements of the lips, jaw, tongue, and face. TD can also involve the limbs, such as repetitive foot tapping and finger movements, and trunk, including rocking and rotatory movements of the neck, shoulders, and hips. It is difficult to rate the severity of TD symptoms from the patient's perspective because even limited symptoms involving just one part of the face can be very noticeable and stigmatizing, and the level of stigma can vary significantly based on the patient's work, school, and/or living situation. Moreover, depending on the underlying disease, patients may not be aware of these involuntary movements, and thus, patient reported symptoms may underestimate their severity. A further challenge in assessing outcomes of treatments for TD is that there are different ways to rate severity of symptoms, including provider's global rating or using a minimum score on the AIMS, calculated either by the treating clinician or by blinded experts reviewing recorded patient videos.

2.3 Treatment Options for Tardive Dyskinesia

Despite over 50 years of experience with antipsychotic drugs and awareness of TD, until recently there were no FDA-approved treatments used to reduce TD symptoms. Many studies have evaluated a wide range of therapies to treat TD, but the overall quality of the studies and the evidence available to support evidence-based recommendations is poor.[13,24] The American Psychiatric Association released a practice guideline for the treatment of patients with schizophrenia in 2004, but it does not specifically address management of antipsychotic drug-related TD.[25] The American Academy of Neurology (see Section 3 for details) published an

evidence-based guideline for treatment of TD, and a systematic review of interventions for treating or preventing antipsychotic drug-induced TD was produced for the U.K. National Institute for Health Research.[13,24] These reports generally highlight an insufficient level of data to make evidence-based recommendations.

Treatment options have included 1) withdrawing the offending antipsychotic drug, 2) changing to a different antipsychotic drug, 3) off-label use of prescription medications, 4) over-the-counter vitamins and supplements, and 5) non-pharmacologic treatments such as electroconvulsive therapy and deep brain stimulation.[2,13,24] Avoiding long-term use of antipsychotics for conditions where evidence of benefit is lacking or other treatment options are available is preferred. However, for many patients with severe psychiatric conditions, antipsychotic drugs may be the best available option. Neither decreasing and then stopping the offending agent nor switching to a different antipsychotic drug has consistently been shown to be an effective treatment option. Moreover, it is often not possible to stop the antipsychotic drug immediately, because TD symptoms can worsen upon withdrawal. Though TD symptoms may improve with these changes, complete resolution of symptoms is uncommon, especially for patients who have taken antipsychotic drugs for long periods of time.[22 26] Long-lasting or permanent symptoms can be seen even in patients who successfully are taken off the antipsychotic drug.[27,28] Therefore, other treatments have been sought to decrease symptoms of patients with TD.

Though a wide range of pharmacologic treatments for TD have been studied, few therapies have been shown to produce more than a slight to moderate benefit.[2,14] Other drugs used have included second-generation antipsychotic drugs, cholinergic agents, anticholinergic agents, antioxidants, benzodiazepines, dopamine agonists, dopamine depleting agents, GABA agonists, calcium channel blockers, buspirone, botulinum toxin and many others.[13,24]

2.4 Vesicular Monoamine Transporter 2 (VMAT2) Inhibitors

The precise mechanism that results in TD from prolonged exposure to DRBAs is not well-understood. One hypothesis involves the secondary upregulation and hypersensitivity of the dopamine receptor.[4,29] Vesicular monoamine transporter-2 (VMAT2) is responsible for storing and releasing dopamine from synaptic vesicles in the brain.[30] The VMAT2 inhibitor tetrabenazine was approved in 2008 by the FDA for use in Huntington's disease. Tetrabenazine has also been used in an off-label manner for individuals with moderate to severe TD symptoms. VMAT2 inhibitors may offset the movement related side effects of antipsychotic drugs and other DRBAs by modulating the pre-synaptic packaging and release of dopamine into the nerve cell synapse. This depletion of dopamine storage in pre-synaptic vesicles results in less dopamine release.

Several small controlled and observational studies of tetrabenazine have shown varying improvement in symptoms, but the need for thrice-daily dosing and side effects, including sedation and worsening of depression and anxiety, have limited its usefulness. In addition, it carries a "black box" warning for depression as well as suicidal thoughts and behavior.

The approval of tetrabenazine for Huntington's disease led to renewed interest in how this mechanism and class of drugs might be used to treat a variety of hyperkinetic movement disorders. Deutetrabenazine (Austedo®, Teva) contains deuterium, a naturally occurring form of hydrogen, which slows metabolism and clearance. Approved for Huntington's disease in April of 2017, it is dosed twice daily and carries the same warnings and contraindications as tetrabenazine. Deutetrabenazine was approved for use in patients with TD on August 30, 2017.

Similar to tetrabenazine and deutetrabenazine, valbenazine (Ingrezza™, Neurocrine Biosciences, Inc.) is a VMAT2 inhibitor, but it has different binding affinities and is dosed once a day. In April 2017 it became the first FDA-approved drug for TD. It does not carry a black box warning or specific contraindications.

Approval by the FDA of the first two treatments for TD represents a potentially important advance for individuals with this frequently irreversible condition. It is noteworthy that the recent systematic review of treatments for TD that was performed before publication of the studies that formed the basis for FDA approval concluded "we still do not know how to treat people with/at risk of TD effectively," and "people with TD feel disappointed and angry at the length of time it has taken for researchers to address the issue of how to treat TD."[13]

Table 2.1. VMAT2 Inhibitors of Interest for the Evidence Review

VMAT2 Inhibitors	Recommended Dose	Dosage Forms & Strengths	FDA Approval Date(S)	Annual WAC (September 2017)
Valbenazine	Initial dose of 40 mg/day. Increase to 80 mg once daily after 1 week	Capsule: 40mg	April 11, 2017 (Tardive dyskinesia)	$75,960*
Deutetrabenazine	Initial dose of 12 mg/day; titrate at a weekly interval by 6mg/day until symptoms improvement, and tolerability, up to a maximum of 48 mg/day given in two divided doses.	Tablets: 6 mg, 9 mg, and 12 mg	April 3,2017 (Huntington's chorea) August 30, 2017 (Tardive dyskinesia)	$90,009†
Tetrabenazine	[Off label use in TD] Recommended dose in Huntington's chorea: Initial dose of 12.5 mg once daily, then titrated by weekly intervals of 12.5 mg/day until symptoms improvement or unacceptable toxicity. Administered in BID if daily dose is less than 37.5mg. If up to 37.5mg to 50mg, administer in three divided doses. If >50 mg/day is required, test and genotype to determine if poor or extensive metabolizers of CYP2D6; not to exceed 100 mg/day or 37.5 mg/dose	Tablets: 12.5mg & 25mg	Aug 15, 2008 (Huntington's chorea)	$19,885¥ – $76,087‡ for 25mg daily dose $79,541¥ - $304,345‡ for 100mg daily dose

*Based on manufacturer input for 80mg dose; †Using the 12mg dose strength for a daily dose of 36mg; ¥ cheapest tetrabenazine generic; ‡Xenazine® cost

2.5 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. Below we summarize the key insights derived from this input.

A recurring theme has been that the common outcome measures used in clinical literature may not adequately capture the impact of TD on social stigmatization, work, family and caregivers, or the treatment of the patient's underlying condition and other co-morbid medical conditions. The involuntary movements caused by TD can cause enormous social stigmatization. For patients already facing the social isolation and symptoms associated with an underlying condition such as schizophrenia or a mood disorder, the added impact of TD can dramatically worsen feelings of isolation, low self-worth, and depression. This can lead to difficulty finding and maintaining a job, and may contribute to lifestyle changes such as inactivity, poor diet, and weight gain that can result in development or exacerbation of chronic medical conditions.

Patients and patient groups identified the lack of valid measures of patient reported symptoms and quality of life with TD as a major limitation of the evidence. Published studies primarily report provider reported outcomes (most commonly the AIMS score). The lack of validated patient-reported outcomes may lead to inaccurate assessment of treatment benefit and potentially underestimate their value. This is made more complex by difficulty in assessing the severity of TD symptoms and by the recognition that some patients may not be fully aware of these symptoms – both leading to uncertainty about how to measure severity and assess the impact of potential treatments. Lastly, there is little evidence that supports translating severity and level of impairment into more formal quality of life measures.

Patients and advocates also raised concern about how TD influences the way clinicians treat the patient's underlying condition. Though physicians may recommend changing to a different antipsychotic drug or stopping such therapy altogether, this often isn't an effective strategy from the perspective of patients. For some, there are no good alternatives to antipsychotic drugs, and changing from one to another may not improve the TD and can adversely affect the control of the underlying condition. Furthermore, stopping all antipsychotic drugs may not lead to decreased TD symptoms, especially for those who have been on these agents for long periods of time. Given the evidentiary and clinical challenges presented by historical attempts to treat TD, patients and patient groups hope that the availability of new, effective treatments for TD may give patients and their clinicians more options for managing both the underlying condition and TD symptoms. There was concern that, because there is currently no data available on how VMAT2 inhibitors will change treatment of the patient's underlying condition, attempts to assess the value may not account for the full impact of these medications on overall quality of life.

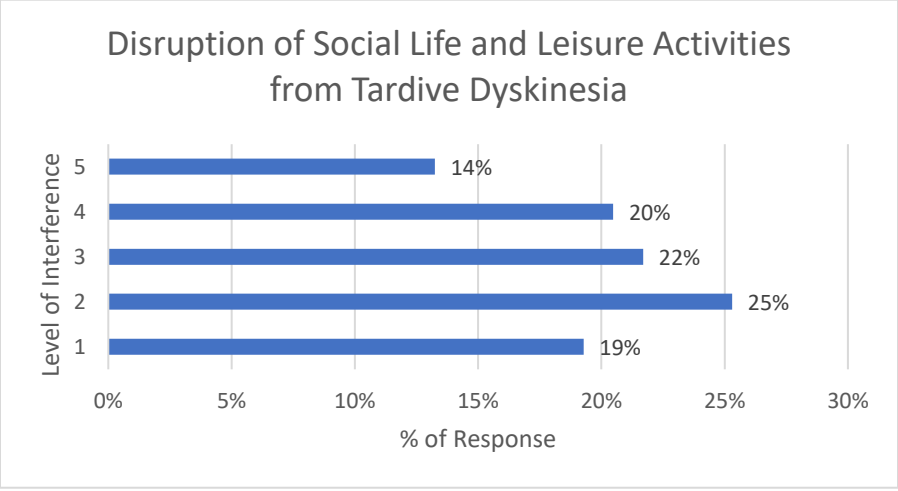
With the advent of the first approved therapies for TD, patient groups wanted to ensure that their members had access to newly available treatment options. Despite limitations previously mentioned, they view these new therapies as giving patients and their clinician more options for treating the underlying condition that required use of DRBAs, and for managing TD symptoms that can dramatically affect the patient, caregiver, and family.

Depression and Bipolar Support Alliance Survey

The Depression and Bipolar Support Alliance (DBSA), a leading patient advocacy group, deployed an online survey to members in August 2017 to gather information about the experience of individuals with TD. A total of 85 individuals responded. Respondents were predominantly diagnosed with bipolar disorder (84%), although patients were also diagnosed with depression (34%), schizophrenia/schizoaffective disorder (15%), insomnia (14%), and stomach problems (3.5%). The duration of TD was generally longstanding: 52% of the sample had symptoms for at least one year, and 25% had symptoms for longer than 6 years. However, 20 patients (24%) reported they no longer had symptoms. Survey respondents reported the many ways doctors attempted to moderate their TD symptoms, including switching the medication that caused the TD to a different one (47%), reducing the dose or stopping the medication that caused the TD (35%), and adding a new prescription or over-the-counter medication to improve the symptoms (40%).

Respondents were asked to describe how much their TD symptoms impacted various aspects of their lives: nearly 60% reported moderate to severe levels of interference (Figure 2.1).

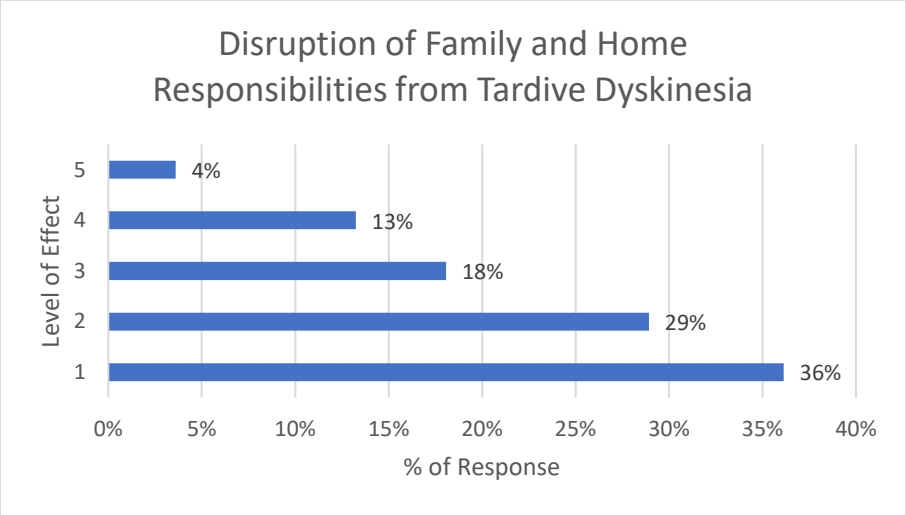
Figure 2.1. How Much Do Health Problems Related to Tardive Dyskinesia Disrupt Your Social Life or Leisure Activities?



1=*No*, 2=*Low*, 3=*Moderate*, 4=*High*, 5=*Severe*

Survey respondents were also asked to rate how much the symptoms of TD disrupt daily responsibilities. In contrast to results for social/leisure impacts, over half of respondents indicated that their symptoms had no effect to minimal effect on daily responsibilities (Figure 2.2).

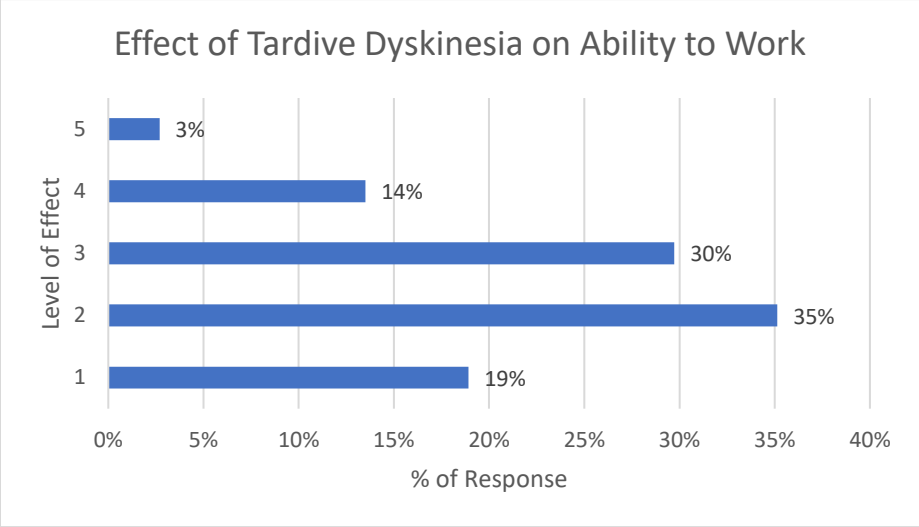
Figure 2.2. How Much Do Symptoms Related to Tardive Dyskinesia Disrupt Your Family Life/Home Responsibilities?



1=*No effect on daily activities, 2=Minimal effect, 3=Moderate effect, 4=Major effect,*
5=*Completely prevented me from doing daily activities*

A total of 37 individuals reported that they were currently employed or had been employed while experiencing TD. Overall, 47% of employed respondents reported a moderate, serious, or completely negative impact on their ability to work (Figure 2.3).

Figure 2.3. How Much Do Health Problems Related to Tardive Dyskinesia Affect Your Ability to Work?



1=No effect, 2=Minimal effect, 3=Moderate effect, 4=Serious effect, 5=Completely prevented me from working

2.6 Definitions of Outcome Measures Used in Clinical Trials

Abnormal Involuntary Movement Scale (AIMS)[23]

The AIMS questionnaire is commonly used to measure the severity of TD symptoms, and the impact of treatment on TD. It is based on the measurement of 12 items, although only seven items (four involving facial/oral movements, two involving the extremities, and one the trunk) were assessed in the clinical trials (see Appendix Table F1). Items 1-7 are measured on a five-point scale of severity from 0-4, with a total score ranging from 0-28. A higher AIMS score reflects increased severity of TD symptoms. Both valbenazine and deutetrabenazine trials used an AIMS score reduction greater than or equal to 50% to define treatment response. Items 8-12 were not measured in any of the clinical trials in our review (8-10 assessed the global severity; and 11-12 assessed problems with the teeth or dentures).

Clinical Global Impression of Change (CGIC)

The CGIC is a clinician-rated measure evaluates the severity of a patient's TD symptoms on a 7-point scale at the time of assessment relative to the clinician's past experience. The score ranges from 1 ("very much improved") to 7 ("very much worse"). In the valbenazine and deutetrabenazine trials patients in the "1" and "2" categories were classified as CGIC responders.

Patients' Global Impression of Change (PGIC)

The PGIC, a 7-point scale reflecting patients' rating of overall improvement, ranges from 1 ("very much improved") to 7 ("very much worse"). In the valbenazine and deutetrabenazine trials patients in the "1" and "2" categories were classified as PGIC responders.

Modified Craniocervical Dystonia Questionnaire (MCDQ-24)[31]

The Craniocervical Dystonia Questionnaire (CDQ-24) is a 24-item patient-reported quality of life assessment tool for use in patients with craniocervical dystonia. It measures five domains: stigma, emotional well-being, pain, activities of daily living, and social/family life. In the trials, certain questions in the five domains were modified to make it relevant to patients living with TD. The modified version (mCDQ-24) was used to assess the impact of TD on the patient's quality of life.

3. Summary of Clinical Guidelines and Coverage Policies

To understand current policy and practice, we identified and reviewed major clinical guidelines for treating TD, and surveyed insurance coverage in New England for pharmacotherapies for TD. We identified one clinical guideline by the American Academy of Neurology (AAN), most recently updated in 2013.[24] All therapies evaluated in the AAN guidelines were for off-label use of prescription drugs, procedures, or over-the-counter vitamins or supplements.

American Academy of Neurology (AAN)[24]

https://www.aan.com/Guidelines/Home/GetGuidelineContent/613

The AAN guidelines found that there was insufficient evidence for many interventions used to treat TD, including withdrawing the DRBA, switching from older (first-generation) to newer (atypical or second-generation) antipsychotic DRBAs, as well as other pharmacologic and health system treatments. The guideline could not recommend for or against withdrawing DBRAs or switching from older to newer antipsychotic DRBAs. Short-term studies suggest that TD symptoms can worsen after withdrawal. Increasing the dosage of or switching to specific older antipsychotic DRBAs may suppress TD symptoms for short periods, but worsen long-term TD, and it may increase the risk of developing akinetic-rigid syndrome. Antipsychotic DRBAs may also cause symptoms of TD and mask rather than treat TD symptoms.

The guideline found limited evidence that tetrabenazine reduces TD symptoms, although there is no long-term evidence, and highlighted a risk of Parkinsonism over the long term. The guidelines suggested that amantadine, clonazepam (in data linked to patients with schizophrenia), and ginkgo biloba may also be effective in reducing TD symptoms; the use of galantamine, eicosapentaenoic acid, and diltiazem was discouraged. Table B1 in the Appendix lists the interventions evaluated in the AAN guidelines.

National Institute for Health Research[13]

https://www.journalslibrary.nihr.ac.uk/hta/hta21430#/abstract

A more recent systematic review sponsored by the U.K. National Institute for Health Research reported similar results. This review found insufficient evidence due to small and low-quality studies, leading to the conclusion that it is not possible to assess the impact of withdrawing DBRAs, changing to new antipsychotic DRBAs or of any specific treatment to reduce the symptoms of TD.[13]

Coverage Policies

We analyzed insurance coverage for AAN-recommended off-label treatment options as well as the recently approved agents evaluated in this review in six New England state Medicaid programs, and 13 silver-tiered insurance plans on individual marketplaces across New England.

All but two commercial plans surveyed covered tetrabenazine. In most circumstances, the generic tetrabenazine was preferred and listed on tiers one or two. Roughly one-third of the plans surveyed required prior authorization for tetrabenazine, and nearly a half of those surveyed required the prescription to be filled at a specialty pharmacy. According to interviews with several payers during our scoping phase, the absolute annual numbers of off-label prescriptions for use of tetrabenazine in TD were very small.

Only three plans covered deutetrabenazine and valbenazine, and each of the three plans listed these therapies on higher level tiers requiring prior authorization. Since both deutetrabenazine and valbenazine were recently approved for their initial indications in April 2017 or later, these lower rates of coverage are unsurprising.

All commercial plans covered both amantadine and clonazepam on their lowest tiers without prior authorization. Ginkgo biloba can be purchased over-the-counter.

In the six New England Medicaid programs, tetrabenazine, valbenazine, and deutetrabenazine were not preferred and all required prior authorization. Amantadine was covered on the preferred drug list in all plans. In Massachusetts and Rhode Island, clonazepam required prior authorization, although it was preferred in the other four states.

Table 3.1. Coverage of Therapies in Major New England Carriers Indicated and Off-Label for Tardive Dyskinesia (See Appendices Table C1 for Full List)

Commercial Plans			
	Covered	Highest Tier(s)	Requires Prior Authorization
Tetrabenazine	85%	27%	31%
Valbenazine	23%	100%	100%
Deutetrabenazine	23%	100%	100%
Amantadine	100%	0%	0%
Clonazepam	100%	0%	0%
New England Medicaid Programs			
	On Preferred Drug List		Requires Prior Authorization
Tetrabenazine	0%		100%
Valbenazine	0%		100%
Deutetrabenazine	0%		100%
Amantadine	100%		0%
Clonazepam	67%		33%

4. Comparative Clinical Effectiveness

4.1 Overview

To inform our analysis of the comparative effectiveness of the VMAT2 inhibitors valbenazine, deutetrabenazine, and tetrabenazine for the treatment of TD, we abstracted evidence from available clinical studies of these agents, whether in published or unpublished form (e.g. conference abstracts or presentations, FDA review documents).

We focused on evidence of the comparative clinical effectiveness of VMAT2 inhibitors in comparison with placebo or other off-label treatments in our target population of adults age 18 years and older with symptoms of TD for at least three months, and history of use of DRBAs. In addition, we examined adverse events and potential harms of VMAT2 inhibitors in populations with conditions other than TD. Importantly, due to key differences in study characteristics, outcomes assessed, and duration of therapy, we did not compare the VMAT2 inhibitors to each other through direct or indirect quantitative assessment.

Our review focused on key clinical benefits common to TD trials including clinician assessed and patient reported outcomes, as well as reported harms.

- Clinical benefits
    - Clinician assessed outcomes including symptom improvement (e.g., Abnormal Involuntary Movement Scale [AIMS], Clinical Global Impression of Change [CGIC])
    - Patient reported outcomes (e.g., Patient Global Impression of Change [PGIC])
    - Health-related quality of life (e.g., modified Craniocervical Dystonia Questionnaire [mCDQ]-24)
- Harms
    - Treatment-related adverse events (e.g. somnolence, headache, worsening of underlying conditions, depression, suicidal ideation)
    - Treatment tolerability (e.g., discontinuation due to adverse events)
    - Death

Where data were available, results for key outcomes were stratified by underlying diagnosis or underlying use of specific DRBAs.

4.2 Methods

Data Sources and Searches

Procedures for the systematic literature review assessing the evidence on VMAT2 inhibitors for TD followed established research methods.[32,33] We conducted the review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[34] The PRISMA guidelines include a checklist of 27 items, further detail of which is available in Appendix Table A1.

We searched MEDLINE, PsycINFO, Cochrane Central Register of Controlled Trials, and EMBASE for relevant studies. We limited each search to English-language studies of human subjects and excluded articles indexed as guidelines, letters, editorials, narrative reviews, case reports, or news items. To supplement the above searches and ensure optimal and complete literature retrieval, we performed a manual check of the references of recent peer-reviewed publications and public reports. Further details of the search algorithms, methods for study selection, and data extraction are available in Appendices A and G.

Study Selection

We included evidence from all relevant published clinical studies irrespective of whether they used a comparative study design. We did not restrict our search by study duration or study setting; however, studies that did not meet a minimum sample size of ten patients were excluded. In recognition of the evolving evidence base for TD, we supplemented our review of published studies with data from conference proceedings, regulatory documents, information submitted by manufacturers, and other grey literature that met ICER standards for review (for more information, see http://icer-review.org/methodology/icers-methods/icer-value-assessmentframework/grey-literature-policy/).

For our review of adverse events, we accepted additional VMAT2 inhibitor RCTs in other movement disorders such as Huntington's disease. Abstracts that reported duplicative data available in published articles were excluded.

Data Synthesis and Statistical Analyses

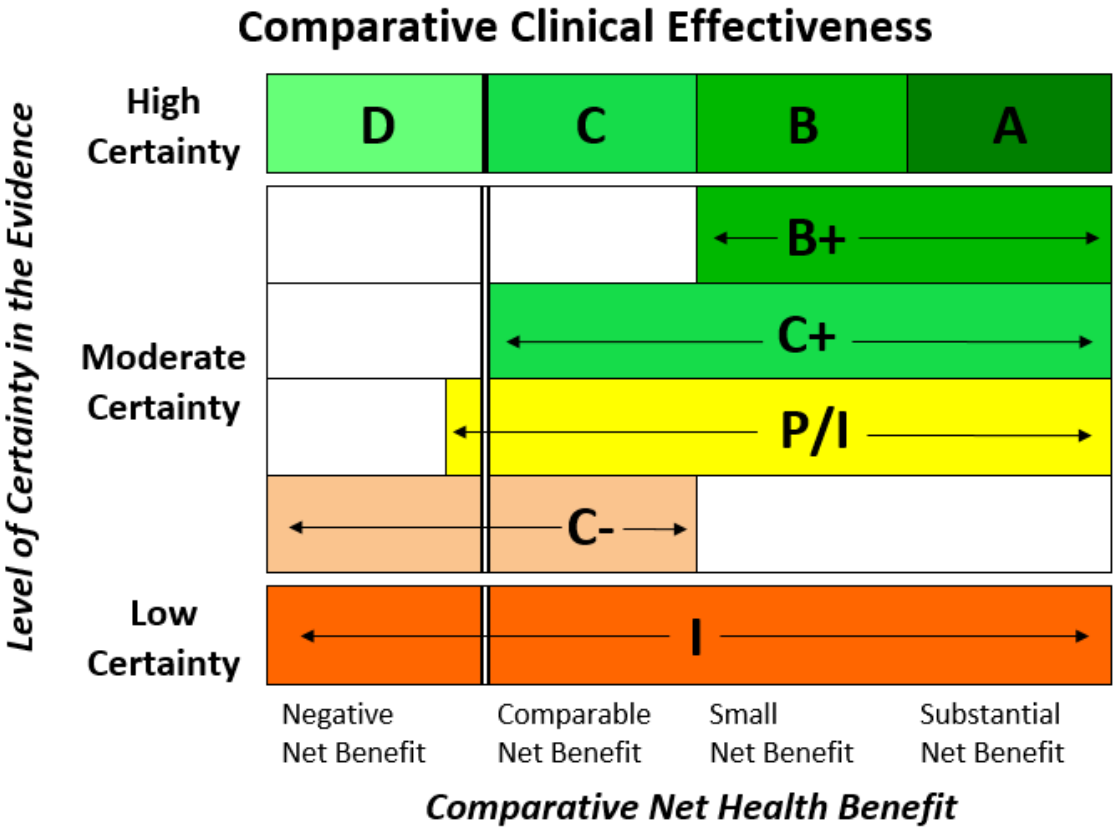
Data on relevant outcomes were summarized in evidence tables (see Appendix G) and are synthesized qualitatively in the text of the report. Due to differences in study characteristics, outcomes assessed, and duration of therapy, we did not compare the VMAT2 inhibitors to each other through quantitative indirect assessment, and therefore focused attention on the comparisons made within the clinical trials of each VMAT2 inhibitor.

Assessment of Level of Certainty in Evidence

We used the ICER Evidence Rating Matrix (see Figure 4.1) 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.[35]

Figure 4.1. 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

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 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 have provided qualitative evidence for use in ascertaining whether there was a biased representation of study results in the published literature. For this review, we identified two completed studies of valbenazine that have not yet been published (NCT01688037 [KINECT trial] & NCT01393600). However, the results of both studies were available in other publicly available sources (FDA medical review document & clinicaltrials.gov) and are discussed as part of the body of evidence.

4.3 Results

Study Selection

Our literature search identified 415 potentially relevant references (see Appendix Figure A1), of which 24 references relating to five comparative trials, one single arm study and five observational studies met our inclusion criteria for assessment. These citations were comprised of 11 publications and 13 conference abstracts/posters. Primary reasons for study exclusion included use of interventions not in our scope, study population outside of our scope (e.g., patients with other hyperkinetic movement disorders), and small sample sizes (n<10). Three additional references were included in our review for the safety assessment. Additional details of included references are described in Appendix F.

Valbenazine

Of the 24 identified references, two publications and six abstracts relating to two placebo-controlled randomized trials (KINECT 2 & KINECT 3), and two open-label extensions (OLE) (KINECT OLE & KINECT 3 OLE) focused exclusively on the use of valbenazine in TD. We included evidence from two additional trials of valbenazine that had no publications or publicly-available conference presentations in our review. Data on one of the trials (NCT01393600) was obtained from clinicaltrials.gov, while the other one (NCT01688037 [KINECT]) was from the FDA medical review document.

Deutetrabenazine

Of the 24 identified references, two publications and seven abstracts relating to two placebo-controlled randomized trials (ARM-TD & AIM-TD), and one OLE (ARM-TD & AIM-TD OLE) focused exclusively on the use of deutetrabenazine in TD. Additionally, we included one reference on the use of deutetrabenazine in Huntington's disease for additional safety assessment.

Tetrabenazine

Of the 24 identified references, seven publications relating to one non-randomized comparative trial, one single arm trial, and five observational studies focused exclusively on use of tetrabenazine in TD. Additionally, we included three references on the use of tetrabenazine in Huntington's disease for additional safety assessment.

Quality of Individual Studies

We rated all four trials of valbenazine and deutetrabenazine to be of good or fair quality using criteria from the U.S. Preventive Services Task Force (USPSTF).[36] One good quality study of valbenazine had comparable study arms at baseline, did not show differential attrition, and used intent-to-treat analysis.[37] The three fair-quality studies also had comparable study arms and did not demonstrate differential attrition, but used a modified intent-to-treat or per-protocol analysis.

Of the two tetrabenazine trials, only one was comparative.[38-40] We judged this trial[41] to be of poor quality due to very small sample size, lack of reliable or valid measurement instrument, and unblinded outcome assessment. All other tetrabenazine studies were single arm studies. Consequently, we did not assign quality ratings to these individual references and instead highlight limitations, uncertainties, and gaps in the evidence in the Controversies and Uncertainties section. We also did not assign a quality rating to references that were obtained from grey literature sources (e.g., conference proceedings).

Comparability of Evidence Across VMAT2 Inhibitors

We did not include tetrabenazine in any quantitative or qualitative comparison involving the other VMAT2 inhibitors due to poor quality data resulting from a lack of controlled trials and other major differences in study designs (e.g., use of nonstandard clinical measures of outcome, very small sample sizes). We did not attempt to conduct a formal comparison of the other two VMAT2 inhibitors (valbenazine & deutetrabenazine) to each other, as the limited number of available studies, as well as differences in trial eligibility criteria, and duration of follow-up precluded these

comparisons (See Table 4.1.). There were also additional differences within the valbenazine and deutetrabenazine trials that are described below (See Table 4.2 & 4.7).

Table 4.1. Comparability of Valbenazine and Deutetrabenazine Trials[a]

	Valbenazine	Deutetrabenazine
Eligibility Criteria	• Moderate to severe TD based on qualitative assessment • Severity based on review of screening videos by multiple external raters	• Moderate to severe TD based on AIMS score ≥6 • Severity criterion required at both screening and baseline assessments • AIMS score assessed by the investigator and confirmed by an independent movement disorder expert via central video rating
Duration of Trials	6 weeks	12 weeks

[a]see Table 4.2 & 4.7 for additional differences within the valbenazine and deutetrabenazine trials

Valbenazine

Data to inform our assessment of the clinical effectiveness of valbenazine (VBZ) were drawn from two published trials of valbenazine: one Phase II trial (KINECT 2 study) and a Phase III trial (KINECT 3 study).[37,38] Both were six week, multicenter, parallel design double-blind, placebo-controlled randomized trials. The six week KINECT 3 study was followed by a 42-week single arm, long-term extension study period.[42]

We identified two other unpublished trials of valbenazine that met our inclusion criteria. One trial was a Phase II parallel design RCT with detailed results available in the FDA medical review document (KINECT study),[43] while the other was a Phase II cross-over design RCT with limited results published on clinicaltrials.gov (NCT01393600). The six week KINECT study was followed by a 6-week single-arm, open-label extension phase. Results from the OLE phase were presented at a clinical conference.[44]

All four RCTs enrolled patients 18 years and older with moderate to severe TD (for at least three months) and history of use of DRBAs for at least three months. Participants were required to have stable psychiatric status, and remain on stable doses of their concomitant medications, including DRBAs, for at least 30 days prior to study entry as well as during the study. There were slight differences in the eligible underlying conditions in each trial. Specifically, the KINECT 3 trial enrolled patients with a clinical diagnosis of schizophrenia/schizoaffective disorder or mood disorder with neuroleptic-induced TD,[37] while KINECT 2 enrolled patients with any DRBA-induced TD including patients with gastrointestinal disorders and metoclopramide-induced TD. Enrollment in the two

unpublished Phase II trials were limited to patients with a clinical diagnosis of schizophrenia/schizoaffective disorder and neuroleptic-induced TD. See Table 4.2 for details of the four studies.

Table 4.2. Valbenazine Trials

	Published Trials		Unpublished Trials	
	KINECT 3	KINECT 2	KINECT	NCT01393600
Study Type	Phase III RCT	Phase II RCT	Phase II RCT	Phase II RCT cross-over design
Total # of Patients	234	102	107	37
Baseline AIMS Score	10	8	15	NR
Mean Age (Years)	56.1	56.2	55.1	51.1
Mean TD Duration	NR	NR	NR	NR
Treatment Duration (Weeks)	6	6	6	Placebo: 2 Valbenazine: 2
Dose/ Day	40mg or 80mg	Starting dose (25mg), then every 2 weeks to max dose of 75mg/d.	50mg or 100mg for 2 weeks & 50mg for 4 weeks	Randomized to 12.5mg or 50mg VBZ after 2 weeks on placebo
Comorbid Psychiatric Condition (%)				
Schizophrenia/Schizoaffective	66.1	58	100	100
Bipolar Disorder/Depression	33.9	38	--	--
Gastrointestinal Disorder	--	4	--	--

NR: Not reported

Clinical Benefits of Valbenazine

The primary efficacy endpoint in the trials of valbenazine was the change from baseline in AIMS score at six weeks. The key secondary outcome was site investigator rated clinical global impression of change (CGIC) mean score. Other secondary outcomes included patient reported global impression of change (PGIC) mean score, and patients classified as responders based on AIMS score (≥50% AIMS improvement), CGIC score ("much improved" & "very much improved"), and PGIC score ("much improved" & "very much improved").

The primary and key secondary endpoints were pre-specified to be analyzed in the intent-to-treat population (KINECT 3) or modified intent-to-treat population (KINECT 2 & KINECT).

Abnormal Involuntary Movement Scale (AIMS)

Published evidence from two trials showed statistically significant improvement in AIMS score for valbenazine compared to placebo. A greater percentage of patients receiving valbenazine achieved a reduction of 50% or more in the AIMS score at six weeks relative to those receiving placebo, resulting in numbers needed to treat of four to seven.

Both KINECT 2 and 3 reported the change in total AIMS score from baseline to week 6 as a primary outcome. The AIMS score was assessed by centrally performed video review and scoring using two blinded raters who were movement disorder specialists not otherwise involved in the KINECT 3 and KINECT 2 trials.[37,38] In contrast, the unpublished KINECT study measured the change using on-site raters who differed from site to site and were not blinded to study visit number.[43]

In the KINECT 3 study (Phase III), the 40mg and 80mg doses of valbenazine resulted in a statistically significant greater reduction in AIMS score compared to placebo at six weeks (least squares [LS] mean change: -1.9 & -3.2 vs. -0.1; $p<0.003$).[37] Similarly, a statistically significant greater proportion of patients who were in the 40mg and 80mg valbenazine groups had an AIMS response (greater or equal to 50% reduction in AIMS score) compared to placebo (23.8% & 40% vs. 8.7%; $p<0.03$).[37] This resulted in a number needed to treat (NNT) of seven and four, respectively, compared with the placebo group. The KINECT 2 trial also showed a statistically significant greater reduction in the AIMS score and higher AIMS response in the valbenazine arm compared to placebo at six weeks (see Table 4.3). Subgroup analysis conducted by psychiatric diagnosis in both the KINECT 2 and 3 trials showed a statistically significant greater reduction in the AIMS score favoring valbenazine over placebo in both the mood disorder and schizophrenia subgroups. This pattern was consistent with the overall population (mood disorder & schizophrenia), although response to valbenazine in the schizophrenia group appears lower than in the mood disorder group (see Appendix Table F2).[45-47]

Table 4.3. Valbenazine: AIMS Change and AIMS Response Across Trials at Six Weeks

Trials	Baseline Score (Mean)	AIMS Reduction from Baseline (LS Mean)	≥50% AIMS Improvement (%)	NNT Vs. Placebo
Kinect 3[37]				
Valbenazine 80mg/D	10.4	-3.2†	40.0†	4
Valbenazine 40mg/D	9.7	-1.9*	23.8*	7
Placebo	9.9	-0.1	8.7	--
Kinect 2[38]				
Valbenazine (Max 75mg/D)	8.0	-2.6†	48.9†	4
Placebo	7.0	-0.2	18.2	--
KINECT‡ (Post-Hoc Analysis)[43]				
Valbenazine (50mg/D + 100mg/50mg)	14.6	-1.2	NR	--
Placebo	15.3	-0.2	NR	--

†p value≤0.001 vs. placebo; *p value≤0.05; ‡unpublished data from FDA Medical Review

The AIMS change observed with valbenazine in the RCT phase of the KINECT 3 trial persisted in the 48 week open-label extension study (Mean change: -3.0 & -4.8 for 40mg & 80mg valbenazine), with 28.3% and 52.4% of patients on 40mg valbenazine and 80mg valbenazine respectively classified as responders at 48 weeks.[42] Importantly, during a four-week washout period at the end of the study, an increase in AIMS score was observed, suggesting a return of TD symptoms following discontinuation of valbenazine.

In contrast to the two published trials, the mean change from baseline in the unpublished Phase II trial (KINECT study) was not significantly different between treatment groups at week 6 (LS mean change: -3.3 vs. -2.5 for valbenazine vs. placebo).[43] As mentioned previously, the primary endpoint in this trial was assessed by on-site raters who were not blinded to the study visit (baseline, two week or four week follow-up visit). A post-hoc analysis of the KINECT trial using blinded central video rating found a greater reduction in the valbenazine group, but this was still not statistically significant (LS Mean change: -1.2 vs. -0.2; p=0.066).[43] In the other unpublished cross over study, the reduction in AIMS score after two weeks on 50mg valbenazine compared to the AIMS score at the end of a two week placebo period did not differ.[48] However, a post-hoc analysis that excluded one of 11 sites (reason for exclusion not provided) showed a significant difference after two weeks on valbenazine compared to the placebo period (Mean difference: -4.2; pvalue=0.0015).[48]

Clinical Global Impression of Change (CGIC)

Evidence is mixed on improvement in TD symptoms as assessed by the CGIC score, with a statistically significant improvement noted for valbenazine over placebo in one trial but no differences observed in two other trials.

The investigator rated CGIC mean score at week 6 was a key secondary outcome in both KINECT 2 and 3. Results from the Phase III KINECT 3 study showed no significant differences between the valbenazine and placebo group on CGIC score and CGIC responders at week 6 in the intent-to-treat population.[37,43] However, an analysis conducted in the KINECT 3 per protocol population (i.e. including only valbenazine treated patients who had a detectable level of valbenazine at week 6) showed significant differences in favor of valbenazine for CGIC score at week 6 (Mean score: 2.8 & 2.8 for 40mg & 80mg valbenazine groups vs. 3.2 for the placebo group; p=0.011).[37] The unpublished Phase II KINECT study also did not show any differences on CGIC after six weeks.[43]

In contrast, the CGIC score in the KINECT 2 study showed a greater improvement in the valbenazine group compared with the placebo group in the modified intent-to-treat population (LS Mean: 2.2 vs. 3.1; p<0.0001).[38] Furthermore, the percentage of patients classified as CGIC responders in KINECT 2 was higher in the valbenazine group, resulting in a NNT of two compared with placebo (Table 4.4).

Although statistical significance was not reported, results of the subgroup analysis conducted by psychiatric diagnosis in the KINECT 2 trial were consistent with the overall population, with a CGIC response of 61.5% vs. 14.8% in the schizophrenia/schizoaffective disorder subgroup, and 73.7% vs. 18.8% in the mood disorder subgroup (valbenazine vs. placebo, respectively).[47]

Table 4.4. Valbenazine: CGIC Score and CGIC Response Across Trials at Six Weeks

Trials/ Arms	CGIC Score (LS Mean)	CGIC Responders (%)	NNT Vs. Placebo
Kinect 3[37]			
Valbenazine 40mg/D	2.9	31.7	NS
Valbenazine 80mg/D	2.9	31.4	NS
Placebo	3.2	20.3	--
Kinect 2[38]			
Valbenazine (Max 75mg/D)	2.2*	66.7*	2
Placebo	3.1	15.9	--
Kinect†[43]			
Valbenazine (50mg/D + 100mg/50mg)	3.3	NR	--
Placebo	3.2	NR	--

*p value≤0.0001 vs. placebo; †unpublished data from FDA Medical Review

Patient Global Impression of Change (PGIC)

Evidence is mixed on improvement in TD symptoms as assessed by the PGIC score, with a statistically significant improvement noted for valbenazine over placebo in one trial but no differences observed in another trial and data indicating inferior performance to placebo in another study.

In the KINECT 2 trial, patients in the valbenazine treatment group had a greater perception of improvement on average, compared with the placebo group (LS Mean: 2.6 vs. 3.3; p=0.0011).[38] The valbenazine group also had a greater proportion of patients classified as PGIC responders compared with the placebo group (57.8% vs. 31.8%, p≤0.001), resulting in a NNT of four.

Although listed as a secondary outcome in the KINECT 3 study, the PGIC outcome was not reported in the published trial results. However, data in the FDA medical review document showed that on average, patients in the valbenazine groups and placebo group did not differ in their report of TD symptom improvement in both the KINECT 3 and KINECT trials.[43] In fact, a significantly *lower* proportion of patients were categorized as PGIC responders in the 80mg valbenazine group compared with the placebo group in the KINECT 3 study. An additional exploratory PGIC responder analysis was conducted to assess the underlying diagnostic categories, and the analysis showed that the difference between 80mg valbenazine group and the placebo group was driven mainly by subjects with schizophrenia or schizoaffective disorder.[43]

Table 4.5. Valbenazine: PGIC Score and PGIC Response Across Trials at Six Weeks

Trials/ Arms	PGIC Score (LS Mean)	PGIC Responders (%)	NNT Vs. Placebo
Kinect 3†[43]			
Valbenazine 40mg/D	2.9	31.7	NS
Valbenazine 80mg/D	3	24.3*	
Placebo	2.7	42.0	--
Kinect 2[38]			
Valbenazine (Max 75mg/D)	2.6*	57.8*	4
Placebo	3.3	31.8	--
Kinect‡[43]			
Valbenazine (50mg/D + 100mg/50mg)	3.2	24.5	--
Placebo	3.1	30.0	--

*p value≤0.001 vs. placebo; †unpublished data from FDA Medical Review

Harms: Valbenazine

The most common side effects of valbenazine were somnolence, fatigue, headache, decreased appetite, akathisia, nausea, vomiting and dry mouth. There was no indication of increased rates of depression and suicidal ideation with valbenazine at 6 weeks.

In the pooled analysis of the KINECT 2 and KINECT 3 trials, the incidence of serious AEs was 6.6% in the valbenazine arm versus 3.9% in the placebo arm. All serious AEs were considered not to be treatment related. There was one death in the valbenazine group, which was considered to be unrelated to treatment.

The most common AE with valbenazine at six weeks was somnolence, occurring in ≥5% of patients and with a higher incidence than in the placebo group.[37,38] Other common AEs with incidence ≥ 3%, and with higher incidence in the valbenazine group compared with placebo group were fatigue, headache, decreased appetite, akathisia (agitation/restlessness), nausea, vomiting and dry mouth. Data from the 48-week long term extension study obtained from a conference abstract a showed similar pattern of common AEs.[45]

Based on similar mechanisms of action and multiple overlapping metabolites of valbenazine and tetrabenazine, and the inclusion of a boxed warning for the risks of depression and suicidality for tetrabenazine's use in Huntington's disease, the FDA's safety review of valbenazine paid particular attention to signals of depression and suicidal ideation.[43] Although, there was one case of suicide attempt in the 6-week KINECT 3 study, findings from both KINECT 2 and KINECT 3 studies showed that the occurrence of depression and suicidal ideation in the valbenazine groups was not significantly different from the placebo-treated patients.[37,38] Furthermore, psychiatric scale assessments indicated no worsening of symptoms in patients with schizophrenia/schizoaffective disorder or mood disorder during long-term valbenazine treatment (see Appendix Table F3-F4).[45]

The pooled discontinuation rate due to adverse events at six weeks from KINECT 2 and KINECT 3 was 4.5% in the valbenazine group, compared with 3.2% in the placebo group (statistical significance not assessed; see Table 4.6). At 48 weeks of follow-up in the KINECT 3 extension study, discontinuation rates due to adverse events in the 40mg and 80mg valbenazine groups were 14% and 18%, respectively.[45]

Table 4.6. Valbenazine: Adverse Events at Six Weeks and 48 Weeks

	Any AE≥1 (%)	SAEs≥1 (%)	Discontinuation Due to AE (%)	Somnolence† (%)	Fatigue (%)	Depression (%)	Suicidal Ideation (%)
6 Week Placebo Controlled Studies (Pooled KINECT 2 & KINECT 3)[37,38]							
VBZ (40mg - 80mg)	45.0	6.6	4.5	6.4	4.0	1.3	2.6
Placebo	39.2	3.9	3.2	4.0	2.4	2.6	5.3
48 Week Long Term Extension Study[45]							
VBZ 40mg	61.9	13.4	13.4	4.0	NR	6.2	5.0
VBZ 80mg	76.2	15.8	17.8	3.1	NR	2.0	5.2

SAE = serious adverse event

Deutetrabenazine

We identified two clinical trials evaluating the effect of deutetrabenazine (DTBZ) on TD: a Phase III trial (AIM-TD) and a Phase II/III trial (ARM-TD). [39,40] Both were 12-week, multicenter, parallel design, double-blind, placebo-controlled randomized trials. The two trials had similar inclusion criteria, enrolling adult patients 18-80 year of age, having an AIMS score ≥ 6 at screening and baseline, a clinical diagnosis of TD forat least three months before screening with stable psychiatric illness, and history of DRBAs use for at least three months (at least one month for patients 60 years of age or older). See Table 4.7 for study characteristics. Both trials were rolled over into the same single-arm, open-label long term extension (OLE) study after a one week washout (54-week data available in conference abstract).[49,50]

Table 4.7. Deutetrabenazine Trials

	AIM-TD	ARM-TD
Study Type	Phase III RCT	Phase II/III RCT
Total # of Patients	293	117
Baseline AIMS Score	NR	9.6
Mean Age (Years)	56.4	54.6
Mean TD Duration (Years)	5.6	6.2
Treatment Duration (Weeks)	12	12
Dose/Day (2 Divided Doses)	Randomized To 12mg, 24mg, or 36mg (Initial Total Dose Of 12mg, Then Titrated Weekly For 4 Weeks Until Randomized Dose	Initial Dose Of 12mg, Then Titrated Weekly For 6 Weeks to Max. Dose Of 48mg/Day or A Significant AE Occurred.
Comorbid Psychiatric Condition (%)†		
Schizophrenia/Schizoaffective Disorder	60	68.4
Bipolar Disorder	17	23.1
Depression	19	25.6
Others	4	--

†The percentages add up to a value greater than 100% in the ARM-TD trial. This is likely because patients could be categorized as having 1 or more comorbid psychiatric condition.

Clinical Benefits of Deutetrabenazine

The primary efficacy endpoint in the trials of deutetrabenazine was the change from baseline in AIMS score. The secondary outcomes were percentage of patients classified as responders based on AIMS score (≥50% AIMS improvement), CGIC score ("much improved" & "very much improved"), and PGIC score ("much improved" & "very much improved"). Additionally, change in modified Craniocervical Dystonia Questionnaire (mCDQ)-24 score was assessed as a secondary endpoint. All outcomes were assessed at 12 weeks.

The efficacy analysis of all outcomes was performed in the modified intent-to-treat population (mITT) who received at least one dose of deutetrabenazine.

Abnormal Involuntary Movement Scale (AIMS)

Published evidence from two trials showed statistically significant improvement in AIMS score for deutetrabenazine compared to placebo. A greater percentage of patients receiving deutetrabenazine achieved a reduction of 50% or more in the AIMS score at 12 weeks, resulting in numbers needed to treat of five compared to placebo.

Table 4.8 presents AIMS score outcomes**.** In the Phase III AIM-TD trial, the 24mg and 36mg doses of deutetrabenazine resulted in a statistically significant greater improvement in the AIMS score compared to placebo at week 12 (LS Mean change: -3.2 & -3.3 vs. -1.4, p<0.003).[39] In addition, significantly more patients in the 24mg and 36mg deutetrabenazine groups had a 50% or more reduction in the AIMS score at 12 weeks compared to the placebo group (35% & 33% vs. 12%; p≤0.05; NNT=5). The ARM-TD trial also showed statistically significant greater improvement in the AIMS score in the deutetrabenazine arm compared to placebo at 12 weeks (see Table 4.8). At 54 weeks in the OLE study, there was further improvement in AIMS score reduction (Mean change: -5.1).[49]

We identified no subgroup analysis by underlying psychiatric condition. Subgroup analysis by ongoing use of DRBAs in the AIM-TD trial showed a statistically significant greater improvement in the AIMS score favoring deutetrabenazine over placebo in both subgroups of patients currently on DRBAs and patients not on DRBA therapy (See Appendix Table F5).[51]

Table 4.8. Deutetrabenazine: AIMS Change and AIMS Response Across Trials at 12 Weeks

Trials	Baseline AIMS Score	LS Mean AIMS Change from Baseline	≥50% AIMS Improvement (%)	NNT Vs. Placebo
AIM-TD[39]				
DTBZ 12mg/D	9.6	-2.1	13	NS
DTBZ 24mg/D	9.4	-3.2†	35†	5
DTBZ 36mg/D	10.1	-3.3*	33†	5
Placebo	9.5	-1.4	12	--
ARM-TD[40]				
DTBZ	9.6	-3.0	--	--
Placebo	9.6	-1.6	--	--

*p value≤0.001 vs. placebo; †p value≤0.05; NS = not significant

Clinical Global Impression of Change (CGIC)

Evidence is mixed on the impact of deutetrabenazine on the CGIC measure, with differences observed in one of three active treatment arms of one trial, and no differences observed in the other trial.

In the AIM-TD study, a statistically significant increase in treatment success on the CGIC was seen in patients receiving deutetrabenazine, but only in the 24 mg/day group (49% vs. 26% for placebo; p=0.014).[39] The 36 mg/day and 12 mg/day groups did not differ from placebo.[39] Similarly, there was no statistically significant difference between the proportion of patients achieving CGIC

treatment success in the deutetrabenazine arm and the placebo arm at week 12 in the ARM-TD trial (see Table 4.9).[40]

An additional post-hoc analysis which in the ARM-TD trial was performed in patients with AIMS ≥6 at both screening and baseline (and included 85.8% of the overall population) resulted in a greater difference between the two arms (52.1% vs 34.7%), although statistical significance was still not reached.[40] However, in a pooled analysis of both AIM-TD (using the AIMS cutoff of the post-hoc analysis) and ARM-TD (24mg/day and 36mg/day groups only), a significantly higher percentage of patients treated with deutetrabenazine achieved treatment success compared with the placebo group (48% vs 30%; *p*=0.005).[52]

Furthermore, in a subgroup analysis of AIM-TD stratified by ongoing DRBA use, CGIC response with deutetrabenazine was statistically better than placebo only in the group not receiving DRBA therapy (See Appendix Table F5).[51]

Table 4.9. Deutetrabenazine: CGIC Treatment Success Across Trials at 12 weeks

Trials	CGIC Treatment Success (%)	NNT Vs. Placebo
AIM-TD[39]		
DTBZ 12mg/Day	28	NS
DTBZ 24mg/Day	49†	5
DTBZ 36mg/Day	44	NS
Placebo	26	--
ARM-TD[40]		
DTBZ	48.2	NS
Placebo	40.4	--
Pooled ARM-TD & AIM-TD[52]		
DTBZ	48†‡	6
Placebo	30‡	--

†p value≤0.05 vs. placebo; ‡unpublished data presented in conference abstract; NS =not significant

Patient Global Impression of Change (PGIC)

Results from two trials showed no significant differences between deutetrabenazine and placebo in the likelihood of achieving treatment success in patients based on the PGIC scale.

In the AIM-TD trial, although there were more patients in the 24mg/ day and 36mg/day deutetrabenazine arms classified as having treatment success compared to the placebo arm, the differences were not statistically significant and patients in the 12mg/day group had a slightly lower response rate than placebo.[39] Similarly, the percentage of patients achieving PGIC treatment success favored the deutetrabenazine arm in the ARM-TD trial, but the difference was not statistically significant (see Table 4.10).[40]

However, in a pooled analysis of data from both the AIM-TD (24mg and 36mg per day groups) and ARM-TD trials, the deutetrabenazine group had a statistically significantly greater proportion of patients with PGIC treatment success compared to the placebo group (43% vs. 30%; p=0.026).[53] Additionally, a subgroup analysis based on the underlying psychiatric condition was performed on the pooled population. The direction of results in both the mood disorder subgroup (deutetrabenazine vs. placebo: 47% vs. 26%) and psychotic disorder subgroup (41% vs. 32%) were consistent with the overall pooled population. However, results were statistically significant only in the mood disorder subgroup (p=0.026).[53]

Table 4.10. Deutetrabenazine: PGIC Treatment Success Across Trials at 12 weeks

Trials	PGIC-TD Treatment Success (%)	NNT vs. Placebo
AIM-TD[39]		
DTBZ 12mg/Day	23	NS
DTBZ 24mg/Day	45	NS
DTBZ 36mg/Day	40	NS
PBO	31	--
ARM-TD[52]		
DTBZ	42.9	NS
PBO	29.8	--
Pooled Analysis (AIM-TD & ARM-TD)[53]		
DTBZ (N=152)	43	8
PBO (N=107)	30	--

NS =not significant

Modified Craniocervical Dystonia Questionnaire (mCDQ-24)

Results showed no significant differences in mCDQ-24 scores in the ARM-TD and AIM-TD trials at 12 weeks.

Although patients receiving 24mg/day and 36mg/day of deutetrabenazine in the AIM-TD trial had greater LS mean reduction from baseline on the mCDQ-24 score compared to placebo, differences were not statistically significant.[39] Patients treated with deutetrabenazine in the ARM-TD trial also had a greater improvement on the mCDQ-24 compared to the placebo group, but these differences were also not significant (See Appendix Table F7). Result of the subgroup analysis from AIM-TD trial was consistent with overall findings.

Harms: Deutetrabenazine

Somnolence, headache, diarrhea, fatigue, insomnia, anxiety, and nasopharyngitis were the most common side effects reported for deutetrabenazine. There was no indication of increased rates of depression and suicidal ideation with deutetrabenazine at 12 weeks.

At 12 weeks, the incidence of adverse events was similar between patients in the deutetrabenazine and placebo groups in both AIM-TD and ARM-TD trials. In the pooled analysis of both trials, the incidence of serious AEs was 6.3% in the deutetrabenazine arm versus 6.9% in the placebo arm. Of the 13 cases of serious AE occurring in the deutetrabenazine group, only one (suicidal ideation) was considered to be possibly related to treatment. There were two deaths in the deutetrabenazine group, and both were not considered to be related to treatment.

Table 4.11 presents the most common AEs and other AEs of interest. At 12 weeks, the most common AE occurring in ≥5% of patients and with higher incidence than in the placebo group were somnolence, headache & diarrhea. Other common AEs that occurred in ≥3% of patients in the deutetrabenazine groups were fatigue, anxiety, nasopharyngitis & insomnia. The incidence of depression and suicidal ideation was similar between treatment groups, and did not show a dose-response relationship. Results from 54-week long-term extension study showed deutetrabenazine was generally well tolerated, with a 5.9% incidence of discontinuation due to adverse events. There was no evidence of increased depression, anxiety, suicidality, restlessness, somnolence and sedation, or Parkinsonism resulting from long-term exposure.

Additional safety information obtained from a 12-week RCT assessing the effect of deutetrabenazine in patients with Huntington's disease (HD) that used a dosing regimen similar to the ARM-TD trial showed a similar side effect profile.[54] The incidence of depression and suicidal ideation was similar between treatment groups in the HD trial (see Table 4.11). However, a generally high rate of depression and suicidality in HD makes it difficult to assess if deutetrabenazine adds to this risk, and because of its chemical similarity to tetrabenazine, a boxed warning for depression and suicidality was added to the deutetrabenazine labelling for HD and continued for its TD indication (see tetrabenazine harms below).[55,56]

Table 4.11. Deutetrabenazine: Adverse Events at Six Weeks and 48 Weeks

	Any AE≥1 (%)	SAEs≥1 (%)	Discontinuation Due to AE (%)	Somnolence† (%)	Headache (%)	Depression (%)	Suicidal Ideation (%)
12 Week Placebo Controlled Studies in TD (Pooled AIM-TD & ARM-TD)							
DTBZ	54	6.3	2.9	5.9	4.9	2	1.5
Placebo	53.4	6.9	3.1	6. 9	7.6	0.8	0.8
54 Week Long Term Extension Study In TD							
DTBZ	59.5	9.5	5.9	7.2	6.9	7.2	1.6
12 Week Placebo Controlled Study In HD							
DTBZ				11.1	NR	4.4	2.2
Placebo				4.4	NR	6.7	2.2

SAE= serious adverse event

Tetrabenazine

We identified seven tetrabenazine studies,[41,57-62] of which only one was comparative.[41] The others were either retrospective or prospective single-arm studies. Tetrabenazine was compared with placebo in a non-randomized cross-over fashion in the comparative study. Unlike the valbenazine and deutetrabenazine trials that required patients to be on a stable DRBA during the study period, patients enrolled in the tetrabenazine trial were required to stop neuroleptic medication at least four weeks before the initiation of tetrabenazine. Furthermore, the tetrabenazine trial was small (n=24), was performed 45 years ago, and used different outcome measures than those used in the valbenazine and deutetrabenazine trials. The majority of the non-comparative trials of tetrabenazine also used a number of different non-standard measures of clinical improvement that make comparing results among these studies problematic.

Clinical Benefit of Tetrabenazine

Tetrabenazine may reduce the symptoms of TD. However, the lack of randomized controlled trials, use of non-standard and variable clinical measures, and other study design limitations (e.g., stopping neuroleptic medications during trial) severely limits any inferences that can be drawn from the available evidence, and precludes any possibility of even qualitative comparisons to the other VMAT2 inhibitors.

In the tetrabenazine cross-over study, 24 patients with clear signs of oral TD were given placebo for four weeks followed by 50mg to 100mg doses of tetrabenazine per day over six weeks in a non-randomized fashion.[41] The goal of the study was to assess the effect of tetrabenazine on TD by

measuring the changes in the frequency of dyskinetic movement in the oral region (e.g. chewing, licking, puffing, pursing of lips, protrusion of tongue) per minute. After six weeks on tetrabenazine, dyskinesia had disappeared completely in eight patients (33%), and reduced significantly in six patients (25%), resulting in a mean oral dyskinesia frequency of 11 per minute in the tetrabenazine group compared to 30 per minute at the end of the placebo period ($p<0.0005$).[41]

We identified only one tetrabenazine study that used the AIMS score as an outcome measure.[59] This was a 20-week, open label single-arm study conducted in 20 patients with TD (mean duration of TD: 40 months) who had failed to respond to previous treatment. Patients were placed on a tetrabenazine dose of 25mg to 150mg per day (mean dose: 57.9mg/day) and were followed prospectively. All patients were required to stop taking the DRBAs four weeks prior to the start of the study, and were videotaped before and after tetrabenazine treatment. At approximately 20 weeks follow up, the mean AIMS score as determined by blinded video raters had improved by 9.7 points from baseline (Mean of 17.9 and 8.2 at baseline and 20 weeks respectively; $p<0.001$).[59]

We identified five other single arm retrospective studies of tetrabenazine conducted in patients with varying hyperkinetic movement disorders including TD where outcomes were separately reported for the underlying disorder (see Appendix Table F8). Response to tetrabenazine therapy was assessed with varying scales of global improvement. In all studies, the scoring was a composite of patient and caregiver assessment, along with the investigator's examination. Results showed a wide range of improvement in the study populations, ranging from 41% to 93% with moderate to marked improvement [61,58] Study descriptions and findings are available in Appendix Table F8.

Harms: Tetrabenazine

The most common side effects of tetrabenazine were somnolence, fatigue, insomnia, fall, depression, agitation, parkinsonism, akathisia, and anxiety. Although statistical significance was not reported, tetrabenazine resulted in substantially greater incidence of depression compared to placebo group in the trial of Huntington's disease.

There were no harms data presented in the comparative study of tetrabenazine for TD. Comparative tetrabenazine safety information was primarily informed by two studies (12 week RCT & 80 week open label study) conducted in patients with Huntington's disease, an indication for which tetrabenazine is approved (see Table F8 in Appendix).[63,64] The first study was a 12-week double blind, placebo controlled trial conducted in 84 patients with Huntington's disease.[63] In this study, tetrabenazine was initiated at a dose of 12.5mg/day and titrated in blinded fashion to a maximum dose of 100mg/day or until the desired antichoreic effect or intolerable adverse effect occurred. The most common AEs occurring in ≥5% of patients and with greater incidence in the tetrabenazine group compared with the placebo group were somnolence, fatigue, insomnia, fall, depression, agitation, and anxiety.[63] Many of the common AEs, including somnolence and

depression, and less common ones, including akathisia and Parkinsonism, were dose limiting, and were managed by dose reductions during the trial.

Notably, the incidence of depression in the tetrabenazine treated patients was substantially greater than in the placebo group (15% vs. 0%).[63] Data from the long term extension study showed that the incidence of depression had risen to 23% at 80-week follow- up.[64] Additionally, there was one case of suicide occurring in a patient with undetected depression during the RCT phase, and one case of a suicide attempt during the long term extension study.[63,64] Based upon the findings from these studies, tetrabenazine was given a black box warning for the risks of depression and suicidality, noting that "tetrabenazine increases the risk of depression and suicidal thoughts and behavior (suicidality) in patients with Huntington's disease. Anyone considering the use of tetrabenazine must balance the risks of depression and suicidality with the clinical need for control of choreiform movements. Close observation of patients for the emergence or worsening of depression, suicidality, or unusual changes in behavior should accompany therapy with tetrabenazine. Tetrabenazine is contraindicated in patients who are actively suicidal, and in patients with untreated or inadequately treated depression."[65]

Review of safety information from the on- and off-label use of tetrabenazine for various hyperkinetic disorders (including TD) in observational studies is summarized in F10 in the Appendix. Adverse events (AEs) that occurred in ≥ 5% of patients in these studies include somnolence, fatigue, parkinsonism, depression, insomnia, nervousness/anxiety, akathisia, and nausea/vomiting.

Controversies and Uncertainties

Ideally, clinical effectiveness of drugs or medical interventions is best informed by evidence from large, high quality, randomized controlled trials. Unfortunately, we did not identify any randomized controlled trials of tetrabenazine for patients with TD. Therefore, our assessment of the use of tetrabenazine in TD was limited and qualitative in nature. We identified four placebo-controlled randomized controlled trials assessing valbenazine and deutetrabenazine, but there were no studies directly comparing these two agents. This is not surprising given that these drugs were approved for use in TD only in April and August of 2017. We also did not identify any trials comparing VMAT2 inhibitors to other agents that have been used to treat TD. As a result, we could only compare the benefits and harms of VMAT2 inhibitors to placebo. Although we considered performing a network meta-analysis in the absence of such head-to-head evidence, the limited number of available studies, as well as major differences in study populations, and duration of treatment precluded such comparisons.

Another major limitation of our evidence base is the lack of outcome measures that directly reflect the burden of TD in affected individuals. Patients and advocacy organizations described the burden of the symptoms of TD on all aspects of a patient's life, including overall quality of life. In addition

to physical impairment, TD can affect employment, interpersonal relationships, and have an impact on other day-to-day activities. The primary outcome in the clinical trials (AIMS scale) was provider-measured, and its measurement was based on office video recordings that may not reflect TD in routine settings. The AIMS scale, though commonly used in research and clinical studies, has well-recognized limitations, including the summation of severity scores across body regions that may not accurately reflect the overall burden of TD symptoms. The need for central, blinded, expert review of video recorded patient assessments highlight the practical challenge in performing such assessments as part of routine care. Moreover, studies demonstrated that assessments done by local investigators who were not blinded to the order of the videos resulted in different outcomes than those from central, blinded reviews.[55]

It is also unclear that the score from AIMS items 1-7 provides an accurate assessment of the overall severity of symptoms. For example, patients with severe facial symptoms that are quite impairing may have the same AIMS score as someone with minimal symptoms in four different body locations, but may feel a much greater burden from facial symptoms because of their place of employment or education. However, none of the trials reported data on the global severity ratings that are included in the AIMS (items 8 -10). Finally, while the statistical significance of changes in AIMS scores were assessed in these trials, there is yet no established threshold for what constitutes a clinically-important difference for change in AIMS score.

Given the impact of TD on quality of life, patient reported measures were included as secondary outcomes in both the valbenazine and deutetrabenazine trials. However, patient reported outcomes comparing valbenazine or deutetrabenazine to placebo generally did not demonstrate a beneficial impact of active treatment, even though TD symptoms were rated as significantly improved by blinded central raters. It is not clear why provider-based and patient-reported measures of success differed. Patients may not have appreciated the amount of benefit that was observed on video reviews. Some experts have suggested that patients with underlying psychotic disorders (e.g., schizophrenia/schizoaffective disorder) may be unaware of their own dyskinetic movements, or their severity, and thus unable to assess the impact of TD on their quality of life. [66,67] Indeed, while results were not definitive, greater improvement in TD symptoms was recorded in some of these studies among patients with mood disorders than in those with schizophrenia. However, as a result of these limitations in measuring the quality of life in patients with TD and how it may change with treatment, there is uncertainty regarding the magnitude of benefit for treatments of TD on patients' overall quality of life. It remains important to develop a patient-based rating scale that can sufficiently capture the individual burden of the disorder in all affected patients.

The side effect profile of the VMAT2 inhibitors is another area of uncertainty. The current clinical trials of valbenazine and deutetrabenazine in patients with TD are limited to only a few weeks'

duration. These trials are not sufficiently long enough to provide the conclusions that their side effect profiles are similar or superior to that of tetrabenazine. However, both valbenazine and deutetrabenazine appear to be well tolerated in the TD clinical trials, despite the addition of a "black box" warning for deutetrabenazine for depression/suicidality (in all likelihood because of its chemical similarity to tetrabenazine). Given the similar mechanism of action and multiple overlapping metabolites with tetrabenazine, the long-term safety of these drugs remains to be determined.

We have heard from patients and their advocates about the concerns of maintaining effective antipsychotic therapy while managing the symptoms of TD. For example, antipsychotic dose reduction by the treating clinician has been used as a strategy to treat or minimize the risk of developing TD and other extrapyramidal symptoms (muscle/posture related side effects of DRBAs including TD). In addition, evidence from the published literature has shown that numerous antipsychotic side effects including TD can lead to non-adherence with therapy.[68] Sub-optimal adherence or deliberate dose-reduction have been shown to increase the risks of psychotic exacerbation and relapse.[69-71] Hence, some stakeholders have suggested that the new VMAT2 inhibitors could have a significant impact on controlling the underlying psychiatric conditions. This is based on the assumption that better TD management will increase adherence to antipsychotics and reduce the use of underlying psychiatric agent modification as a treatment strategy. While this is a plausible concept, it has not been evaluated in clinical studies to date, and so real-world data will be needed to assess these effects.

Finally, while the development of the new VMAT2 inhibitors is a step forward in the management of TD, there are still several unanswered questions regarding the use of this class of drugs in clinical practice. Future studies should evaluate the patient or clinical factors (e.g. age, TD symptoms, baseline AIMS score, TD duration, ongoing use of DRBAs) that predict the likelihood of response to VMAT2 inhibitors. In addition, pragmatic clinical trials that include patients with other comorbid conditions or severe mental disorders encountered in routine clinical practice will be helpful. Furthermore, future studies should evaluate not only the short-term effect of these drugs but also the factors that will predict the durability of these effects and long term functional improvement.

Summary

We reviewed data on use of VMAT2 inhibitors for the treatment of adults aged 18 and older with symptoms of TD for at least three months and history of use of DRBAs. Using the ICER Evidence Matrix, we assigned evidence ratings independently for each of the VMAT2 inhibitors compared to placebo.

Table 4.12. ICER Evidence Ratings

VMAT2 Inhibitors	ICER Evidence Rating
Valbenazine	P/I
Deutetrabenazine	P/I
Tetrabenazine	I

Valbenazine

- Treatment with valbenazine resulted in a greater reduction in AIMS scores and more patients with a substantial improvement in AIMS scores compared to placebo.
- Despite improvements in AIMS scores, treatment with valbenazine did not consistently result in improvement on Clinical Global Impression of Change (CGIC) scores. Additionally, currently available data on patient reported outcome (PGIC scores) do not demonstrate a benefit for valbenazine over placebo.
- Although valbenazine was generally well tolerated in the six-week trials, this is a new therapy that requires ongoing use, and important adverse effects could become apparent over time in larger patient populations.

For adults with TD for at least three months and history of use of DRBAs, we have moderate certainty that valbenazine provides a comparable, small, or substantial net health benefit. However, because of the lack of clear benefit on CGIC scores and the patient reported PGIC scores, as well as the absence of long-term safety data, we cannot definitively rule out the possibility of A SMALL net harm. Therefore, we consider the evidence on valbenazine to be "promising but inconclusive" (P/I).

Deutetrabenazine

- Treatment with deutetrabenazine resulted in a greater reduction in AIMS scores, and more patients with a substantial improvement in AIMS scale compared to placebo.
- Despite improvements on AIMS score, treatment with deutetrabenazine did not result in improvement on CGIC scores. Additionally, currently available data on patient reported outcome (PGIC scores) do not demonstrate a benefit for deutetrabenazine over placebo.

- Although deutetrabenazine was generally well tolerated in the 12-week trials, this is a new therapy with a black box FDA warning for depression and suicidality that requires ongoing use, and important adverse effects could become apparent over time.

For adults with TD for at least three months and history of use of DRBAs, we have moderate certainty that deutetrabenazine provides a comparable, small, or substantial net health benefit. However, because of the lack of clear benefit on the clinician assessed CGIC scores and the patient reported PGIC scores, as well as the absence of long-term safety data, we cannot definitively rule out the possibility of a small net harm. Therefore, we consider the evidence on deutetrabenazine to be "promising but inconclusive" (P/I).

Tetrabenazine

- No high-quality comparative studies were identified on the use of tetrabenazine in adults with TD and history of use of DRBAs.
- Currently available evidence on tetrabenazine suggest a possible benefit for the use of tetrabenazine in improving symptoms of TD; however, the use of various non-standard clinical measures assessed over variable periods of time and without concurrent comparators makes interpretation of results extremely problematic.
- The tolerability profile of tetrabenazine was reported in a clinical trial of patients with Huntington's disease. Using doses similar to that in TD, tetrabenazine resulted in more patients with adverse events (e.g., somnolence and insomnia), including some serious adverse events (e.g., depression) compared to placebo. However, many of these events were dose related and managed by dose adjustments without discontinuing tetrabenazine.

Although tetrabenazine appears to have some potential benefit, we cannot be certain whether tetrabenazine is comparable, or possibly even inferior to placebo or other therapies due to the lack of direct comparative evidence in patients with TD and the potential unfavorable safety profile. We therefore consider the evidence for tetrabenazine in adults with TD for at least three months and history of use of DRBAs to be "insufficient" (I).

5. Long-Term Cost-Effectiveness

5.1 Overview

We conducted a lifetime cost-effectiveness analysis using a simulation model of the two FDA-approved VMAT2 inhibitors, valbenazine and deutetrabenazine, for treating the symptoms of moderate-to-severe TD compared to placebo in adult patients with underlying schizophrenia/schizoaffective disorders and bipolar and major depressive disorders (hereafter referred to as "other affective disorders"). As noted in Section 4.3, differences in trial design, entry criteria, and outcome measurement precluded explicit comparisons between these two agents, and so the cost-effectiveness of each was separately compared to placebo.

While other drug and non-drug treatments are used off-label for managing TD symptoms, these were not included in our analyses, as the quality of evidence supporting their use in TD was poor.[13] Other treatment strategies, such as altering the dose of the medication thought to be causing TD symptoms, also were not included because of similar concerns of low-quality or insufficient evidence, and because such changes were not allowed in the study protocols for valbenazine and deutetrabenazine. We included model parameters such as the probability of TD symptom improvement, utility gain from symptom improvement, treatment discontinuation rates, mortality rates, treatment costs, and other health care costs. We incorporated clinical parameters based on applicable clinical trial data. Utilities, mortality rates, and treatment costs were obtained from drug manufacturers and published literature identified through the systematic review. Health care costs related to TD treatment were unavailable in the literature and were estimated based on expert opinion.

The primary outcomes of the model included discounted total payer costs, life years, quality-adjusted life years (QALYs) gained, and incremental cost-effectiveness ratios, using a health-care system perspective over a lifetime horizon. A scenario analysis using a modified societal perspective including productivity costs was also conducted.

5.2 Cost-Effectiveness Model: Methods

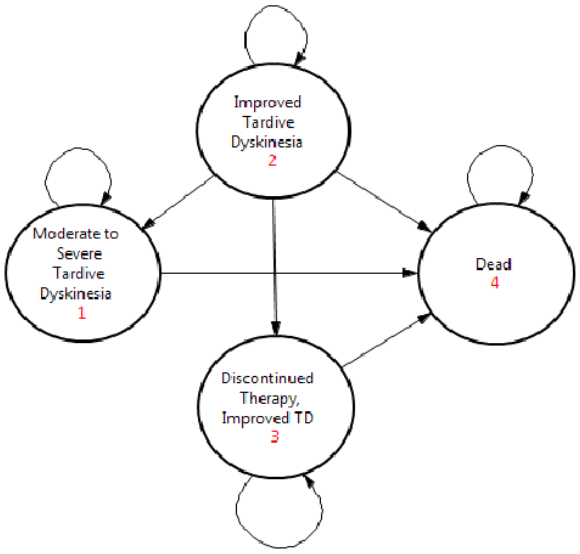
Model Structure

We developed a new semi-Markov model with time-dependent mortality and discontinuation rates for this evaluation. There were four health states in the base-case model for the treatment arms: improved TD (where patients remained on treatment for TD), moderate to severe TD (where

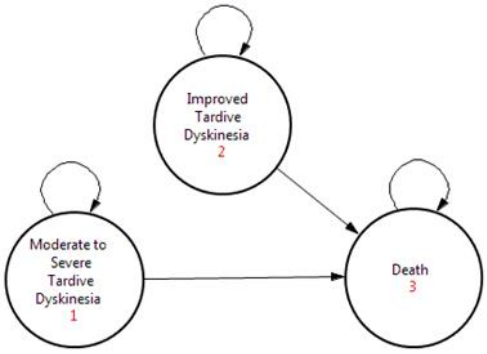
patients had discontinued TD therapy, either after an initial one-month trial or at a later time due to adverse events or lack of efficacy), discontinued therapy with improved TD (patients had improved TD symptoms despite discontinuing therapy), and death (Figure 5.1). Because there is no discontinuation of treatment in the placebo arm, three health states were used to define the model as follows: moderate to severeTD, improved TD, and death (Figure 5.1). Note that, because treatment effects were observed rapidly in the clinical trials, they are modelled ahead of the start of the annual cycles in the model; because of this, patients cannot move from moderate to severe TD to improved TD after the model has begun. For the base case, patient responses to valbenazine and deutetrabenazine were defined as 50% improvement in the Abnormal Involuntary Movement Scale (AIMS) scores observed in clinical trials. The models were built using Microsoft Excel 2013 (Redmond, WA).

Figure 5.1. Base Case Model Structures

Treated Patients



Placebo Patients



Key Model Characteristics

The base case model used a U.S. health care system perspective (i.e., focused only on direct medical care costs), and employed annual cycles over a lifetime horizon.

Prior to entering the first cycle of the model, all patients in the treatment arm were modelled as receiving one month of treatment. Patients initiating therapy would have a trial period of the medications resulting in either response and continuation of treatment or non-response and

discontinuation, after which they entered into the model and cycled through the four health states based on annual transition probabilities, until all but 0.1% of patients had transitioned to death. For the placebo model, patients entered the model based on placebo response rates. Mortality was modeled using age-adjusted mortality rates from CDC/NCHS National Vital Statistics System. In addition, the model incorporated the "life-table" method of half-cycle adjustment, which is similar to the original half-cycle correction method but more robust.[72] Costs and quality adjusted life years were discounted at an annual rate of 3%, beginning with the second model cycle.

Target Population

A review of the literature was conducted, separate from the clinical review described above, to identify model inputs for demographic characteristics of patients with TD and patients with schizophrenia, schizoaffective, and affective disorders taking medications known to cause TD. From the identified papers, those best representing the intended modeled population were selected for the model inputs.

The underlying proportion of patients receiving antipsychotic medications were obtained from Domino et al, which characterized new users of antipsychotic medications using the Medical Expenditure Panel Survey (MEPS) from 1996 to 2005. In the base case model, the proportion of patients with schizophrenia/schizoaffective disorders was set at 70.2% and the proportion with bipolar and other affective disorders was set at 29.8%.

Patient age and gender was necessary to estimate mortality risk for the modeled population. Data was obtained from two separate studies characterizing the population aged 18-64 years with schizophrenia/schizoaffective disorder[73] and affective disorders.[74] In these studies, patients with schizophrenia/schizoaffective disorders had a mean age of 38 years and were 52.5% female[73] while patients with affective disorders had a mean age of 40 years and were 64.8% female.[74] Using these population estimates resulted in a modeled population with an expected average age of 39 years.

The resulting modeled population represented U.S. adults ages 18-64 years with TD and the underlying conditions of schizophrenia/schizoaffective disorders or affective disorders, and on medications that cause TD (as a proxy for patients with moderate to severe TD).

Key Model Assumptions

Table 5.1 contains a list of key model assumptions along with the rationale for each assumption.

Table 5.1. Key Model Assumptions and Rationale

Assumptions	Rationale
Patients receiving metoclopramide for gastrointestinal conditions and developing tardive dyskinesia were not modeled in this analysis.	Patients receiving metoclopramide were excluded from most of the primary trials involving valbenazine or deutetrabenazine. Furthermore, there is insufficient data regarding the demographics of patients receiving metoclopramide and developing TD to accurately depict this population in the model.
Before entering the model, all patients received one month of treatment and incurred one month of treatment costs as an initial cost.	Patients initiating therapy would have a trial period of the medications. Medications are usually dispensed in one-month increments.
Patient response to the first month of treatment was reflected in their initial health states.	This prevented a lag in response to therapy that would occur if all patients were started in the "moderate to severe tardive dyskinesia" state. Additionally, patient response to therapy was quite rapid in the clinical trials.[36,75]
Patients not responding to treatment with valbenazine or deutetrabenazine discontinued their treatment.	Patients not responding will not see benefits of therapy and should have no TD treatment costs.
Response to treatment remained constant for all responders. Patients did not improve or decline beyond their initial response to therapy while remaining in the "improved TD" state.	There is limited information on the individual change in response to therapy over time. Furthermore, there is no information available on the impact of TD severity on quality of life.
Long-term discontinuation rates were modeled from open-labeled studies with less than one year of observation.[36,76] Following the first cycle, discontinuation rates were modeled as being 50% of that observed in the first cycle.	There is no information regarding discontinuation rates of therapies beyond the clinical trial extensions. We lowered the discontinuation rate by 50% following the first year because for most therapies, patients typically discontinue their medications at a higher rate in the first year of treatment.
Patients who responded to treatment were assumed to have no added primary care and neurologist visit costs related to TD. Patients not responding to treatment were assumed to have two additional primary care and two additional neurologist visits per year.	There is currently no data on the costs associated with treating TD. It is likely that patients whoseTD has improved will incur fewer office visits.
TD and treatments do not have a direct effect on mortality.	Neither TD nor treatment with valbenazine or deutetrabenazine have been demonstrated as having an impact on mortality.
Treatment of TD has no effect on the outcomes or costs of treating the underlying conditions.	We identified no studies that demonstrate valbenazine or deutetrabenazine improve the management of underlying conditions.

Interventions

Interventions assessed in this model included the two FDA-approved VMAT2 inhibitors, valbenazine and deutetrabenazine, both compared with placebo. Patients not receiving active treatment for TD symptoms were modeled using outcomes associated with the placebo arms from the relevant clinical trials.[36,75] We modeled the effects and costs of the highest doses reported in clinical trials (valbenazine 80mg and deutetrabenazine 36mg per day) because those doses were generally associated with the highest effects. All patients were assumed to have received one month of therapy prior to entering the model.

Initial State Probabilities

Patients entered the model as having improved or not, following the first month of therapy. Those who improved on treatment entered the model in the "improved TD" state. Those who had not improved began the model in the "moderate to severe TD" state.[36,75] The proportion of patients starting in the "improved TD" state were abstracted from the Phase III clinical trials evaluating response to valbenazine and deutetrabenazine compared with placebo (Table 5.2). The effectiveness of valbenazine and deutetrabenazine was based on the proportion of patients that experienced a ≥50% improvement in the AIMS score at the maximal tolerated dose in the clinical trials (80mg for valbenazine and 36mg for deutetrabenazine). The ≥50% improvement was chosen because it was consistently reported for both therapies, and as a balance between the available utility gains in the literature based on "full improvement" and the goals of the clinical trials.

Table 5.2. Key Model Inputs: Probabilities

Parameters	Valbenazine		Deutetrabenazine	
	Value	Reference	Value	Reference
Proportion of PLACEBO Responders with ≥50% Reduction in AIMS	8.7%	Hauser 2017[36]	12.0%	Jimenez-Shahed 2017[75]
Proportion of TREATMENT Responders with ≥50% Reduction in AIMS	40.0%	Hauser 2017[36]	33.1%	Jimenez-Shahed 2017[75]
Annual Discontinuation Rate (First Year)	19.0%	Remington 2016[77] Hauser 2017[37]	13.0%	Anderson 2017[78]
Annual Discontinuation Rate (Year 2+)	9.5%	Calculation*	6.5%	Calculation*

*Assumed 50% decrease after 1st year

Transition Probabilities

Adverse Events and Medication Discontinuation

In clinical trials, serious adverse events were uncommon and occurred at a rate similar to what was observed in the placebo arm. Hence the model did not directly incorporate the impact of adverse events in either arm. The model did incorporate discontinuation rates, however. As discussed above, the initial state included responders and non-responders, and non-responders are assumed to discontinue their TD treatment permanently and start out in the "moderate to severe TD" state. Following the initial state, discontinuation rates in the treatment arm are applied to the responders who start in the "improved TD" state, based on the longest follow-up data available corresponding to the first year within the model. Specifically, the discontinuation probability for the first year of valbenazine was calculated based on the discontinuation rate from the longest reported observation period from open label studies, subtracting the discontinuation rate from the clinical trial, and then extrapolating to one year.[36,77] The discontinuation probability for the first year of

deutetrabenazine was equal to the exposure-adjusted incidence rate reported in Anderson et al., which only included discontinuations after a washout period.[76] For both drugs, we assumed a 50% decrease in discontinuation rate after the first year for all subsequent years (Table 5.1).

Among initial responders who subsequently discontinued, a portion of patients entered the "discontinued therapy, improved TD" state and the remainder entered the "moderate to severe TD" state, where it was assumed that those responding to treatment would include all those that would have been placebo responders in the absence of treatment and that those who would not have been placebo responders would return to the moderate to severe TD state upon treatment discontinuation. For example, if all patients discontinued their TD therapy, then the resulting proportion in the moderate to severe TD state would be the same as the placebo response rate. Therefore, patients discontinuing therapy and continuing to have improved TD with no treatment was calculated using the following conditional probability:

(Proportion with placebo response) / (Proportion of treatment responders).

The remainder of patients that discontinued therapy cycled back into the moderate to severe TD state.

Mortality

Mortality was modeled using age- and gender-adjusted mortality estimates for the general population. Mortality estimates were obtained from the CDC/NCHS National Vital Statistics System, and further adjusted to reflect mortality rates in the underlying populations of schizophrenia/schizoaffective disorders and other affective disorders. The multipliers used to adjust population mortality rates were derived from the published literature, and were 3.70 and 2.05 for schizophrenia/schizoaffective and other affective disorders, respectively (Table 5.1).[79,80] We assumed that TD and TD-related therapies did not affect mortality.[13]

Cost Inputs

Costs associated with caring for TD were included in the model (Appendix Table H1). Specifically, medication costs and primary care/specialist visits to manage TD were considered. Annual drug acquisition costs for valbenazine (80mg per day) and deutetrabenazine (36mg per day) were used in the model. Since both drugs were only recently approved by the FDA, information on their net sales volume and revenue was not available. For deutetrabenzine, we applied the industry-wide average discount rate of 27% for branded drugs[81] to the wholesale acquisition cost (WAC) of the 12 mg tablets to arrive at an annual net price of approximately $64,000. The manufacturer reported an expected annual list price for the 80mg capsule of valbenazine of $75,960, to which we applied a discount of 27% to derive an annual net price at approximately $55,500.

Patients with moderate to severe TD were assumed to require additional neurologist and primary care visit costs. Specifically, having moderate to severe TD was assumed to necessitate two additional primary care and two additional neurologist visits per year. Costs for these visits were derived from the 2017 Medicare Physician Final Fee Schedule (approximately $365) (Appendix Table H3).[82] There were no additional costs associated with death.

Utilities

Gains in utility from improvement in TD symptoms were modeled using utility estimates available in the literature for moderate to severe TD versus no TD (Appendix Table H4). The mean utility for the modeled population with improved TD was 0.82, which reflects utility scores associated with the underlying conditions.[83,84] To this utility, we applied a utility decrement of 0.095 to those patients with moderate to severe TD.[85] We chose the utility decrement for TD because it was based on standard gamble utilities from subjects rating TD relative to perfect health and was directly elicited from healthy patients upon viewing the symptoms of moderate to severe TD, independent of any underlying conditions. Patients in the "improved TD" or "discontinued therapy, improved TD" states did not incur the disutility. The utility gained from improvement in TD was assumed to be independent of any other underlying condition(s). There is limited data on the impact of TD on utility and no data by severity level. We used an estimate that is conservative, in that it allocates a benefit associated with the complete removal of TD to patients with a 50% or higher decrease in the AIMS score, biasing the model results in favor of the therapies.

Sensitivity Analyses

One-Way Sensitivity Analyses

All model parameters were subjected to one-way sensitivity analysis. In addition to one-way sensitivity analyses, threshold analyses were conducted varying the discontinuation rate or the annual acquisition costs of valbenazine and deutetrabenazine until (when possible) there was a projected incremental cost effectiveness ratio of $150,000/QALY gained.

Probabilistic Sensitivity Analysis

A Monte Carlo probabilistic sensitivity analysis was conducted using 10,000 simulations. Distributions were assigned based on available data. Beta distributions were assigned for probabilities and for disutility associated with TD.[86] Parameters for specifying beta distributions were obtained from literature values using the number of occurrences of the event in question and the overall sample size. The disutility of TD symptoms also used a beta distribution with a mean score equal to the base case, to provide a relatively large variance that was also bounded between zero and one. Distributions describing drug costs and the medical costs of managing TD were

unavailable. Given this absence of information, uniform distributions were used, with costs being varied between 50% and 150% of the base case values. All probabilistic sensitivity analysis distributions and parameters are shown in Appendix Table H5.

Scenario Analyses

Because utility measures evaluating different severity levels of TD were not available, and there was considerable uncertainty surrounding the impact of VMAT2 inhibitors on utilities, results were also presented using TD symptom-reduced years (i.e., proportion of patients with a 50% reduction in TD symptoms each model cycle). This measure encompassed the additional average number of years individuals could expect to live with reduced symptoms of TD while on therapy compared with placebo. To produce these estimates, all patients with improved TD received a value of 1 and all patients with moderate to severe TD received a value of 0. The model was then run and the cost per TD symptom-reduced year calculated.

In addition to varying base-case model parameters, an analysis from a modified societal perspective was undertaken, varying the annual lost productivity due to TD from $0 to $2,252. The $2,252 value was based on employment differences observed in a single study that included patients with schizophrenia and TD.[87] In this study, 22.9% of patients without TD were employed, while 17.8% with TD were employed. The median 2016 US salary ($44,148) was used to estimate an upper bound for productivity benefits resulting from improved TD [calculated as a 5.1% difference in proportion of patients who worked and median annual US income (i.e. 0.05 * $44,148 = $2,252)].

Currently, there is little agreement on the best measure for assessing patient TD symptoms or improvement. In addition to the AIMS, clinical trials involving valbenazine and deutetrabenazine included the clinician-rated CGIC scale. A beneficial response to therapy was stated as being either a ≤ 2-point improvement in the valbenazine trials or a response of "improved" or "very much improved" in the deutetrabenzine trials (See Section 2.5). Scenario analyses were conducted evaluating the impact of VMAT2 inhibitors on the proportion of responders when using the CGIC scale. Results were presented as cost/QALY gained and as cost per TD symptom-reduced year.

Finally, scenario analyses were also undertaken to account for potential indirect effects of valbenazine and deutetrabenazine on the underlying conditions of the patients. Separate models were developed for patients with schizophrenia/schizoaffective disorders and bipolar disorders. The models incorporated well-controlled and poorly-controlled mental health states for the underlying conditions, to reflect a proportion of patients with TD becoming non-adherent to their antipsychotic medication. For this scenario, it was assumed that 10% of patients with symptomatic TD (in non-responders to treatments or placebo) and well-controlled mental health for the underlying condition would become poorly controlled for one cycle. This was intended as a rough approximation of the potential impact of non-adherence related to TD symptoms. Further, as there is no data for this potential impact, the

probability of having poor control was also varied between 0% and 100% in this scenario analysis. Further details on the structure, health states, assumptions, and additional parameter inputs for this analysis are available in Appendix H.

5.3 Cost-Effectiveness Model: Results

Base-Case Results

The main results are summarized in Tables 5.3-5.5. The total discounted lifetime costs for valbenazine and the placebo comparator were approximately $176,000 and $6,900, respectively (Table 5.3). The total discounted quality-adjusted life years (QALYs) for valbenazine and placebo were 15.34 and 15.12, respectively. Deutetrabenazine and its placebo comparator had lifetime discounted costs of approximately $220,000 and $6,600 and lifetime discounted QALYs of 15.37 and 15.18, respectively. The incremental cost-effectiveness ratios over a lifetime horizon were approximately $750,000 per QALY for valbenazine and approximately $1.1 million per QALY for deutetrabenazine (Table 5.4).

Table 5.3. Base-Case Discounted Costs and Outcomes

Treatment Arm	Total Costs	Total QALYs
Valbenazine	$176,235	15.34
Placebo (Valbenazine Comparison)	$6,876	15.12
Deutetrabenazine	$220,277	15.37
Placebo (Deutetrabenazine Comparison)	$6,627	15.18

Table 5.4. Pairwise Results for VMAT2 Inhibitors Compared to Placebo

Regimen	Incremental Costs	Incremental QALYs	Incremental Cost-Effectiveness Ratios vs. Placebo
Valbenazine	$169,359	0.22	$754,440
Deutetrabenazine	$213,650	0.19	$1,100,025

Costs per TD symptom-reduced year were approximately $72,000 and $105,000 for valbenazine and deutetrabenazine versus placebo respectively (Table 5.5). It is difficult to judge the importance of these results, however, as there are no clear benchmarks of cost per TD symptom-reduced year for comparison.

Table 5.5. Cost per TD Symptom-Reduced Year

Treatment	Cost per TD Symptom-Reduced Year
Valbenazine vs. Placebo	$71,672
Deutetrabenazine vs. Placebo	$104,502

Sensitivity Analyses over Different Time Horizons

The estimated incremental cost-effectiveness ratios were higher over shorter time horizons and all surpassed commonly utilized thresholds of $100,000 or $150,000 per QALY in the base case analysis. The full results for these sensitivity analyses are reported in Appendix H5.

One-Way Sensitivity Analyses

Results of the one-way sensitivity analyses are shown in tables 5.6 and 5.7. Altering the proportion of responders in either treatment or placebo group across calculated 95% confidence intervals did not result in an incremental cost effectiveness ratio below the $150,000 per QALY gained threshold. Similarly, one-way changes in discontinuation rates, drug costs, cost of TD, and disutility across reasonable bounds resulted in incremental cost effectiveness ratios that were consistently above the $150,000 per QALY gained threshold.

Figure 5.2. Tornado Diagram for Valbenazine One-Way Sensitivity Analyses

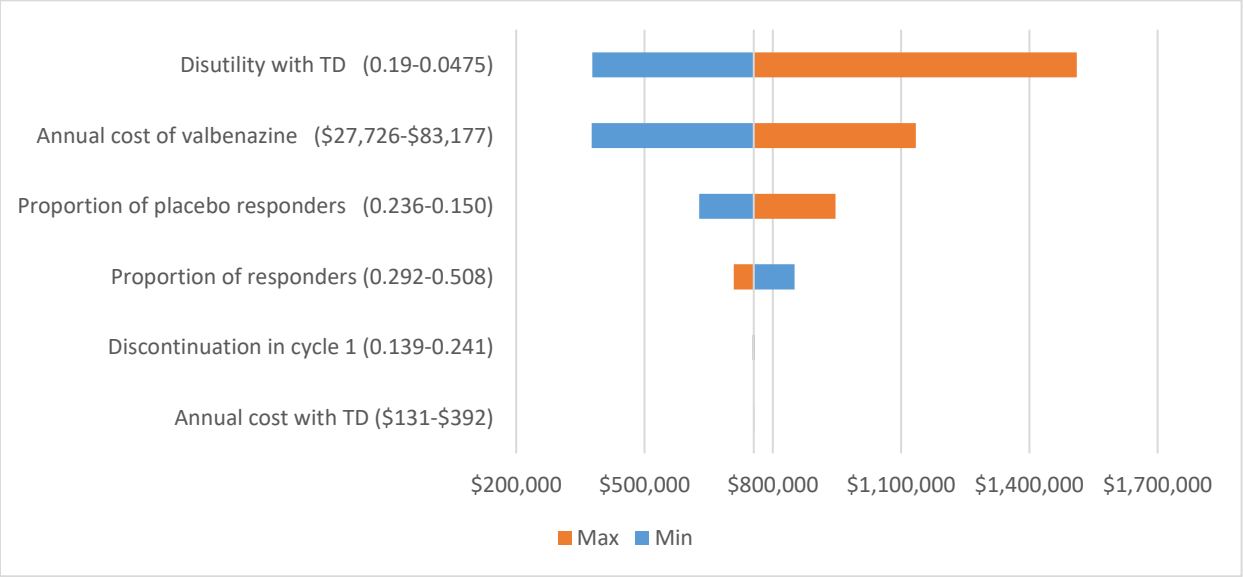


Table 5.6. One-Way Sensitivity Analyses: Valbenazine

Parameters	Base Case Value	Sensitivity Analysis Value		Incremental Cost per QALY	
		Min	*Max*	*at input min*	*at input max*
Proportion of Responders: Placebo	0.087	0.0236	0.150	$627,716	$945,528
Proportion of Responders: Valbenazine (80mg)	0.40	0.292	0.508	$849,815	$707,999
Annual Discontinuation *	0.19	0.139	0.241	$752,111	$756,823
Annual Cost of Valbenazine	$55,451	$27,726	$83,177	$375,297	$1,133,584
Annual Cost of TD Care	$261.54	$131	$392	$756,363	$752,518
Utility Decrement due to TD	-0.095	-0.0475	-0.19	$1,508,880	$377,220

*Note that the model assumes annual discontinuation rates in cycles 2 and beyond are 50% of first-year rates

Figure 5.3. Tornado Diagram for Deutetrabenazine One-Way Sensitivity Analyses

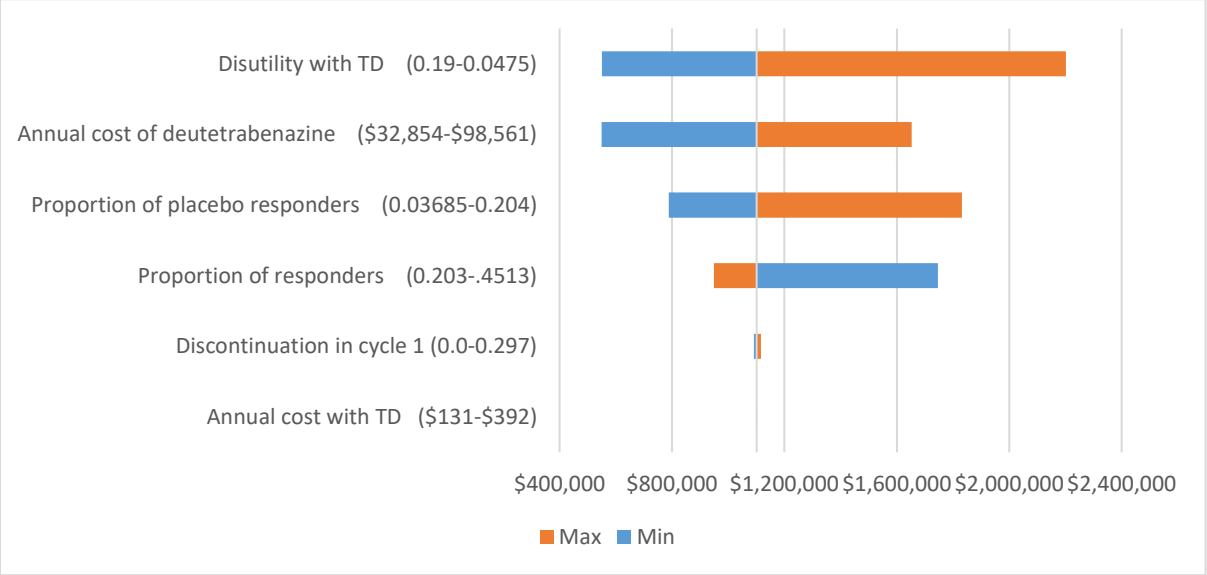


Table 5.7. One-Way Sensitivity Analyses: Deutetrabenazine

Parameters	Base Case Value	Sensitivity Analysis Value		Incremental Cost per QALY	
		min	max	at input min	at input max
Proportion of Responders: Placebo	0.12	0.03685	0.204	$787,984	$1,830,143
Proportion of Responders: Deutetrabenazine	0.331	0.203	0.4513	$1,744,907	$948,182
Annual Discontinuation*	0.13	0.0	0.297	$1,090,029	$1,115,562
Annual Cost of Deutetrabenazine	$65,707	$32,854	$98,561	$548,089	$1,651,960
Annual Cost of TD Care	$261.54	$131	$392	$1,101,947	$1,098,102
Utility Decrement due to TD	-0.095	-0.0475	-0.19	$2,200,049	$550,012

*Note that the model assumes annual discontinuation rates in cycles 2 and beyond are 50% of first-year rates

Lower discontinuation rates and higher treatment effects resulted in lower incremental cost effectiveness ratios. However, neither lowering annual discontinuation probabilities to 0% nor increasing the proportion of treatment responders to 100% in one-way sensitivity analyses resulted in incremental cost effectiveness ratios below the $150,000 per QALY threshold. Utility gains from treatment would need to reach 0.48 for valbenazine and 0.70 for deutetrabenazine (base case = 0.09) to reach a cost-effectiveness threshold of $150,000 per QALY gained.

Probabilistic Sensitivity Analysis

The results of the probabilistic sensitivity analyses are summarized in Appendix H5. At willingness to pay thresholds of $150,000 per QALY or lower, zero percent of the simulations reach these thresholds for valbenazine or deutetrabenazine relative to placebo.

Threshold Analysis

In single-input threshold analyses, values below $150,000 per QALY gained were observed when the following values were used for inputs (Table 5.8).

Table 5.8. Threshold Analysis Results for Valbenazine and Deutetrabenazine

	WAC Per Unit	WAC Per Month	Unit Price To Achieve $50,000 Per QALY	Monthly Price To Achieve $100,000 Per QALY	Unit Price To Achieve $100,000 Per QALY	Discount From WAC To Reach Thresholds
Valbenazine 80mg per Day		$75,960	$3,936	$7,596	$11,256	85%-95%
Deutetrabenazine 36mg per Day		$90,009	$3,204	$6,180	$9,156	90%-96%

Scenario Analyses

Including productivity gains from improved TD in the model, with all else remaining equal, resulted in only small improvements in the cost-effectiveness ratios for the two agents (see Appendix Figure H3 and Table H6), but these finding did not approach commonly-cited thresholds for cost-effectiveness. Two additional scenario analyses attempted to account for potential but unproven benefits of TD treatment in improved maintenance and medication adherence for patients' underlying conditions of schizophrenia/schizoaffective disorders and bipolar disorders. The incremental cost effectiveness ratios for these scenario analyses were predictably lower compared with the base case analysis (Table 5.9), but also remained well above commonly-cited thresholds. In patients with the underlying condition of schizophrenia/ schizoaffective disorders, the estimated incremental cost effectiveness ratios were $555,000 per QALY gained for valbenazine and $779,000 per QALY gained for deutetrabenazine. For patients with bipolar disorder, the incremental cost effectiveness ratios for valbenazine and deutetrabenazine were $607,000 and $874,000 per QALY gained, respectively. Varying the probability of non-adherence resulting in poorly-controlled schizophrenia or bipolar disorder to 100% (a highly unlikely scenario) resulted in cost-effectiveness ratios that still exceeded $150,000 per QALY (see Appendix).

Table 5.9. Scenario Analyses of Control Benefit from Improved Adherence

Base Case Vs. Placebo	Underlying Condition	Incremental Cost Per Qaly
Valbenazine	Schizophrenia/Schizoaffective Disorders	$554,932
Deutetrabenazine	Schizophrenia/Schizoaffective Disorders	$778,734
Valbenazine	Bipolar Disorder	$606,971
Deutetrabenazine	Bipolar Disorder	$874,315

Model Validation and Prior Published Evidence on Costs and Cost-Effectiveness

Model validation followed standard practices in the field. We tested all mathematical functions in the model to ensure they were consistent with the report (and supplemental Appendix materials). We also conducted sensitivity analyses with null input values to ensure the model produced findings consistent with expectations. Three independent modelers tested the mathematical functions in the model as well as the therapy-specific inputs and corresponding outputs.

We also compared the ICER model to previously published models. We searched the literature to identify models that were similar to our own, with comparable populations, settings, perspective, and treatments. In our review of the literature, we found only one model that was relevant as a comparison to our model.

A model was developed by Rosenheck to assess the cost-effectiveness of second-generation antipsychotics in reducing the risk of TD in adults with schizophrenia.[88] Reported incremental cost-effectiveness ratios at one year ranged from approximately $660,000 per QALY gained using any one of four second-generation antipsychotics relative to perphenazine, a first-generation antipsychotic, to approximately $1.7 million per QALY gained when using risperidone relative to perphenazine, a first-generation antipsychotic. Extending the model time-horizon to five years for the same interventions and comparator decreased the incremental results to approximately $264,000 and $684,000 per QALY gained, respectively. The interventions, comparators, associated costs, and effectiveness in reducing TD were different between the ICER and Rosenheck models; therefore results cannot directly be compared. However, comparison from a structural perspective revealed other key differences between the two models: 1) a life-time horizon in the ICER model versus a maximum five-year time horizon in the Rosenheck model; 2) 3% discounting applied to costs and consequences in the ICER model versus no discounting in the Rosenheck model; and 3) population with approximately 70% with schizophrenia in the ICER model versus only patients with schizophrenia in Rosenheck. We also note that the Rosenheck model employed a different measure to assess improvement in TD symptoms.

5.4 Potential Budget Impact

We used the cost-effectiveness model to estimate the potential total budgetary impact of valbenazine and deutetrabenazine for patients with TD resulting from use of antipsychotics. We used the wholesale acquisition cost (WAC), an estimate of discounted WAC, and the three cost-effectiveness threshold prices for each drug in our estimates of budget impact.

Potential Budget Impact Model: 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 a new therapy rather than 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 primary 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.

The potential budget impact analysis included the entire candidate population for treatment, which consisted of adult patients with TD resulting from use of antipsychotic drugs used to treat underlying disease. Dhir et al. estimated the prevalence of TD using US population estimates from 2016 with epidemiological data on the prevalence and use of antipsychotic drugs to treat underlying schizophrenia, bipolar disease and major depressive disorder (MDD).[89] The estimated prevalence of TD in 2016 was 573,000 patients in the US, of which 359,000 had moderate-to-severe TD. The authors also reported projected total prevalence and prevalence by severity of TD up to 2025. For our budget impact analysis, we used the projected prevalence of moderate-to-severe TD between 2017 (359,000 patients) and 2021 (361,000 patients) to arrive at an average prevalence estimate of 360,000 patients over five years, or 72,000 patients each year.

ICER's methods for estimating potential budget impact and calculating value-based benchmark prices are described in detail elsewhere and have recently been updated. The intent of our revised approach to budgetary impact is to document the percentage of patients that could be treated at selected prices without crossing a budget impact threshold that is aligned with overall growth in the US economy.

Briefly, we evaluate a new drug that would take market share from one or more drugs, and calculate the blended budget impact associated with displacing use of existing therapies with the new intervention. In this analysis, we assumed that each of the VMAT2 inhibitors would only replace the cost of disease manangement with non-specific therapy since there have been no FDA-approved therapies for TD prior to the approval of the two reviewed VMAT2 inhibitors.

Using this approach to estimate potential budget impact, we then compared our estimates to a budget impact threshold that represents a potential trigger for policy mechanisms to improve affordability, such as changes to pricing, payment, or patient eligibility. As described in ICER's methods presentation, this threshold is based on an underlying assumption that health care costs should not grow much faster than growth in the overall national economy. From this foundational assumption, our potential budget impact threshold is derived using an estimate of growth in US gross domestic product (GDP) +1%, the average number of new drug approvals by the FDA each year, and the contribution of spending on retail and facility-based drugs to total health care spending. Calculations are performed as shown in Table 5.10.

For 2017-18, therefore, the five-year annualized potential budget impact threshold that should trigger policy actions to manage affordability is calculated to total approximately $915 million per year for new drugs.

Table 5.10. Calculation of Potential Budget Impact Threshold

Item	Parameter	Estimate	Source
1	Growth in US GDP, 2017 (est.) +1%	3.20%	World Bank, 2016
2	Total health care spending, 2016 ($)	$2.71 trillion	CMS NHE, 2014
3	Contribution of drug spending to total health care spending (%)	17.7%	CMS National Health Expenditures (NHE), 2016; Altarum Institute, 2014
4	Contribution of drug spending to total health care spending ($) (Row 2 x Row 3)	$479 billion	Calculation
5	Annual threshold for net health care cost growth for ALL new drugs (Row 1 x Row 4)	$15.3 billion	Calculation
6	Average annual number of new molecular entity approvals, 2013-2014	33.5	FDA, 2016
7	Annual threshold for average cost growth per individual new molecular entity (Row 5 ÷ Row 6)	$457.5 million	Calculation
8	Annual threshold for estimated potential budget impact for each individual new molecular entity (doubling of Row 7)	$915 million	Calculation

Potential Budget Impact Model: Results

Table 5.10 and 5.11 illustrate the per-patient budget impact calculations in more detail, based on WAC, discounted WAC and the prices to reach $150,000, $100,000, and $50,000 per QALY for valbenazine and deutetrabenazine compared to usual care. Note that annual estimated budget impact is substantially lower than the annual WAC for both drugs given rates of non-reponse and discontinuation seen in clinical trials.

Table 5.11. Per-Patient Budget Impact Calculations for Valbenazine Over a Five-year Time Horizon

Average Annual Per Patient Budget Impact					
	WAC	Discounted WAC	$150,000/QALY	$100,000/QALY	$50,000/QALY
Valbenazine	$30,349	$22,219	$4,252	$3,226	$1,773
Usual Care	$345				
Difference	$30,003	$21,874	$3,940	$2,881	$1,427

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

Table 5.12. Per-Patient Budget Impact Calculations for Deutetrabenazine Over a Five-year Time Horizon

Average Annual Per Patient Budget Impact					
	WAC	Discounted WAC	$150,000/QALY	$100,000/QALY	$50,000/QALY
Deutetrabenazine	$31,286	$22,908	$3,405	$2,375	$1,345
Usual Care	$333				
Difference	$30,953	$22,576	$3,073	$2,043	$1,013

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

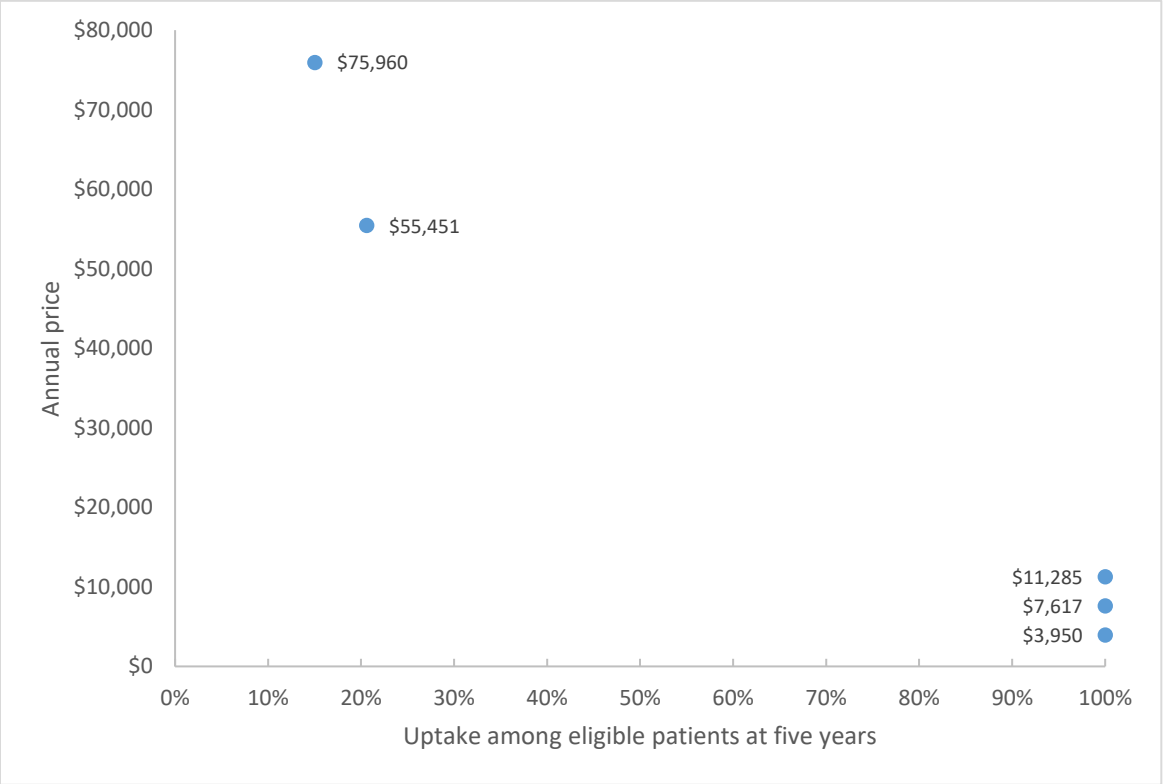
Valbenazine

The average potential budgetary impact when using the annual WAC ($75,960) was an additional per-patient cost of approximately $30,000, and approximately $21,900 using the annual discounted WAC ($55,451). Average potential budgetary impact at the three cost-effectiveness threshold prices for the drug ranged from approximately $3,900 per patient using the annual price ($11,285) to achieve $150,000 per QALY to approximately $1,400 using the annual price ($3,950) to achieve a $50,000 per QALY cost-effectiveness threshold.

As shown in Figure 5.4, approximately 15% of eligible patients could be treated in a given year without crossing the ICER budget impact threshold of $915 million at WAC ($75,960) and approximately 21% of patients at the discounted WAC ($55,451). The entire elgible cohort could be

treated without crossing the ICER annual budget impact threshold of $915 million at the three threshold prices.

Figure 5.4. Valbenazine Pricing and Uptake Relative to Annual Budget Impact Threshold

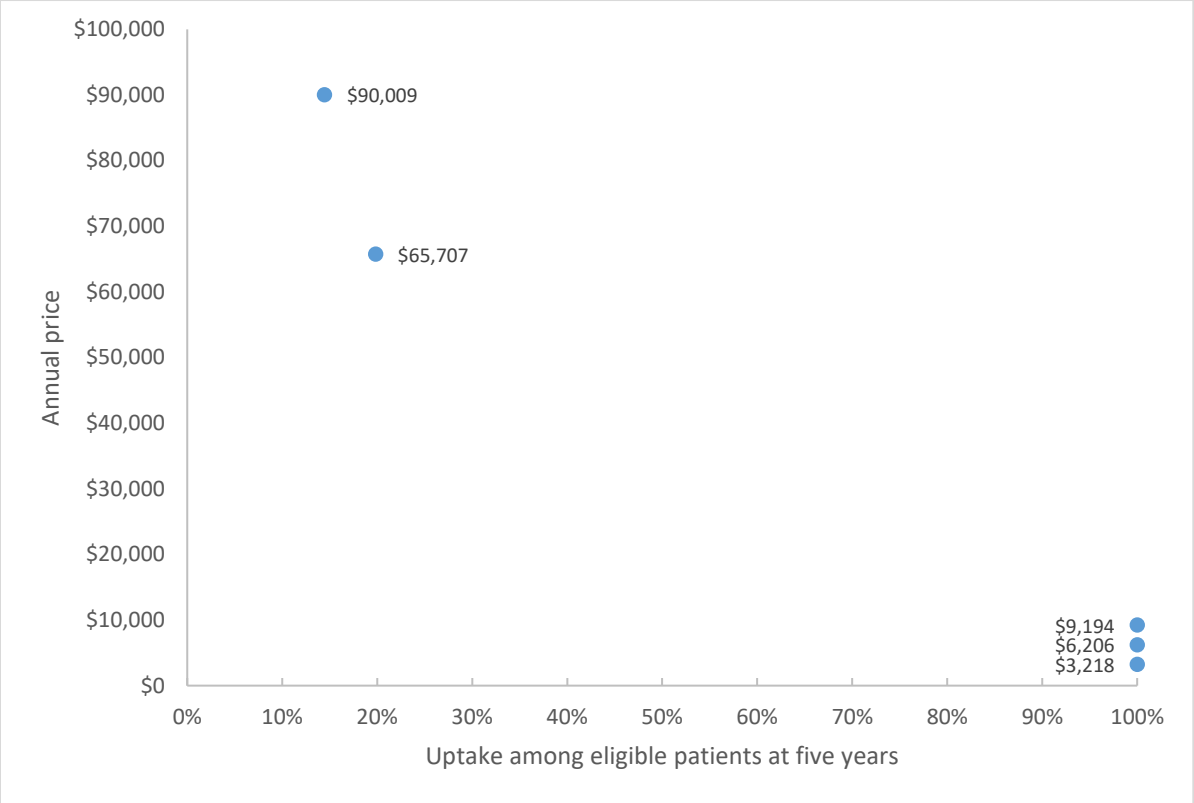


Deutetrabenazine

The average potential budgetary impact when using the annual WAC ($90,009) was an additional per-patient cost of approximately $31,000, and approximately $22,600 using the annual discounted WAC ($65,707). Average potential budgetary impact at the three cost-effectiveness threshold prices for the drug ranged from approximately $3,100 per patient using the annual price ($9,194) to achieve $150,000 per QALY to approximately $1,000 using the annual price ($3,218) to achieve a $50,000 per QALY cost-effectiveness threshold.

As shown in Figure 5.5, approximately 14% of eligible patients could be treated in a given year without crossing the ICER budget impact threshold of $915 million at WAC ($90,009) and approximately 20% of patients at the discounted WAC ($65,707). The entire elgible cohort could be treated without crossing the ICER annual budget impact threshold of $915 million at the three threshold prices.

Figure 5.5. Deutetrabenazine Pricing and Uptake Relative to Annual Budget Impact Threshold



5.5 Value-Based Benchmark Prices

Value-based benchmark prices will be released in the revised Evidence Report, which will be posted on November 21, 2017.

5.6 Summary and Comment

We estimated the cost-effectiveness of the two recently-approved VMAT2 inhibitors compared to placebo in patients with TD. In the base case, the incremental cost effectiveness ratios for valbenazine and deutetrabenazine versus placebo far exceeded commonly utilized cost-effectiveness thresholds. When model inputs were varied between reasonable ranges in one-way sensitivity analyses, none of the resulting estimates approached thresholds of $150,000 per QALY. Further, the probabilistic sensitivity analyses resulted in acceptability curves suggesting an extremely low likelihood that the treatments will reach these thresholds.

There were several limitations of this analysis. First, the effectiveness of the TD treatments was based on limited intermediate measures from the clinical trials. Furthermore, current measures for assessing TD severity may not accurately reflect the global impact of TD on patients' lives or the impact of TD severity on overall quality of life. As a result, the impact of improvement from the treatment of TD on utility was not available in the literature. We attempted to overcome this limitation by generously applying the reported difference in utility between moderate-to-severe TD and no TD for responders in our analysis, based on a 50% reduction in the AIMS score. In addition, the model applies to an average TD patient, but may not reflect particular sub-populations or individuals. However, our reporting is consistent with guidelines for assessing and reporting cost-effectiveness analyses and, as noted in Section 4, findings for treatment effect were generally consistent across subgroups identified in the clinical trials.

Another limitation to this model was a lack of data on discontinuation of TD medication due to adverse events beyond the first year. We acknowledge that this input influences model results and hence had to make assumptions on post-year one discontinuation rates. Finally, due to a paucity of information available in the literature, an important limitation was that expert opinion was used for estimating the annual non-drug medical costs of treating TD. However, the model was insensitive to this input, requiring an input of over 250 times the base case value (of ~$262) to achieve commonly-cited cost-effectiveness thresholds.

Finally, our budget impact analyses for the use of VMAT2 inhibitors in place of placebo (i.e., no TD-focused treatment) indicated that, when using estimated net prices, only 21% and 20% of the entire eligible TD cohort could be treated with valbenazine and deutetrabenazine respectively before crossing the budget threshold of $915 million per year.

In summary, our analyses suggest that the clinical benefits associated with valbenazine and deutetrabenazine will lead to increased quality-adjusted life expectancy over no specific treatment for TD symptoms (i.e., placebo). At current pricing levels, however, the estimate lifetime cost-effectiveness of these agents far exceeds commonly-cited cost-effectiveness thresholds.

6. Other Benefits and Contextual Considerations

Our reviews seek to provide information on other benefits offered by VMAT2 inhibitors to the individual patient, caregivers, the delivery system, or the public that would not have been considered as part of the evidence on comparative clinical effectiveness. These elements are listed in the table below.

Table 6.1. Potential Other Benefits and Contextual Considerations

Potential Other Benefits
This intervention provides significant direct patient health benefits that are not adequately captured by the QALY.
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 who have failed other available treatments.
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 surveillance with no maintenance therapy, there is significant uncertainty about the long-term risk of serious side effects of this intervention.
Compared to surveillance with no maintenance therapy, 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.

The review of available comparative clinical evidence on the use of VMAT2 inhibitors for TD due to DRBAs highlights key gaps in knowledge that may limit the ability to assess the potential benefit of these drugs. For example, because of the challenges of capturing the social impact of TD on individual patients, patients and patient groups expressed concern that the outcome measures

reported in clinical literature may underestimate the magnitude of the impact of these drugs. We heard from several patients about the significant impact TD can have on ability to work. Although data are lacking on the effectiveness of VMAT2 inhibitors on work-related outcomes, there is reason to believe that for some patients with TD, VMAT2 inhibitors may make it easier to find a job and/or maintain a job.

Furthermore, some stakeholders highlighted that VMAT2 inhibitors may have a positive impact on adherence to treatment of the underlying psychiatric and other medical conditions, especially if DRBAs are needed. Although there are no data available on these potential benefits, it was noted that for patients with serious psychiatric conditions in whom DRBAs are the most effective treatments available, patient concerns about side effects, including TD, may influence their decision to take DRBAs or comply with therapy if prescribed. More effective treatment for individuals with TD in whom DRBAs are the therapy of choice may improve compliance with DRBA therapy and the control of the underlying psychiatric condition.

Among the VMAT2 inhibitors, tetrabenazine requires three times per day dosing, deutetrabenazine requires twice a day dosing, and valbenazine requires once daily dosing. Adherence with treatment may be somewhat improved by using medications that require less frequent dosing, although the degree of potential improvement is unclear. However, potential differences in side effects with long-term use are more likely to be important for adherence than dosing frequency.

For many patients, TD is a chronic condition. Since data suggests stopping VMAT2 inhibitors leads to a recurrence of TD symptoms, treatment will require long-term use. As newly approved drugs, the long-term risks of valbenazine and deutetrabenazine will only become apparent with ongoing use in a large number of treated individuals. Currently, the only comparative data for VMAT2 inhibitors is to placebo, and the duration of the randomized portion of clinical trials was only 6-12 weeks. Thus, it is unclear how valbenazine and deutetrabenazine may compare to each other and to other therapies that have been used off-label for TD in terms of effectiveness and safety with longer-term use. In addition, the magnitude and durability of ongoing use of VMAT2 inhibitors is uncertain and will require prospective study.

Although many questions remain about the benefits and harms of VMAT2 inhibitors for TD, there is considerable interest among patients and their healthcare professionals in having an FDA approved treatment for this condition. Approval of these drugs for TD represents a potentially important advancement for individuals with this frequently irreversible condition. In particular, for individuals with severe disabling TD, the use of VMAT2 inhibitors may decrease caregiver/family burden.

Potential Cost-Saving Measures in Tardive Dyskinesia

We sought to identify areas of waste and low-value care in psychiatry that could be reduced to make way in health care budgets for new innovations. We reached out to clinicians, patients and patient groups, manufacturers, and other payers for input on potential targets for waste reduction. The following areas were highlighted by stakeholders:

- *Stakeholders from Mental Health America (MHA)* suggested that early intervention programs can be cost-saving over the long term for standard mental health services by delivering better initial intervention for patients with first-time psychosis. The suggestion is strengthened by a small study conducted in Australia which found cost savings in a small group of patients who were treated for up to two years with an early psychosis intervention, and yielded better health outcomes and lower health care costs eight years later.[90]
- *Stakeholders from the Depression and Bipolar Support Alliance (DBSA)* highlighted the results from a 2015 DBSA survey on agitation and emergency care, which showed that about 60% of patients presenting with agitation in the ED returned within a year (34% within 60 days). They suggested that patients should be routinely provided with information about mental health services in the community or be referred to psychiatrists upon discharge from the ED to decrease return visits and their associated costs to the health care system.

Additionally, we reviewed the American Board of Internal Medicine's Choosing Wisely® campaign, which encourages specialty societies to identify areas of low-value care that could be reduced or eliminated. Below are recommendations we identified that could potentially be cost-saving.

- *Choosing Wisely* suggests ensuring an appropriate initial evaluation and ongoing monitoring for patients before prescribing antipsychotic medications for any indication due to metabolic, neuromuscular, and cardiovascular side effects. [91] These side effects may potentially result in additional costs.
- *Choosing Wisely* also recommends against prescribing antipsychotics as first choice to treat behavioral and psychological symptoms of dementia, and as a first-line intervention for insomnia in adults, because the evidence demonstrates that the risks (e.g. cerebrovascular effects, parkinsonism, extrapyramidal signs, TD, confusion, increased body weight) outweigh the potential benefits in these populations.[91,92]

References

1. Li P, Snyder GL, Vanover KE. Dopamine Targeting Drugs for the Treatment of Schizophrenia: Past, Present and Future. *Current topics in medicinal chemistry.* 2016;16(29):3385-3403.
2. Cloud LJ, Zutshi D, Factor SA. Tardive dyskinesia: therapeutic options for an increasingly common disorder. *Neurotherapeutics.* 2014;11(1):166-176.
3. Tarsy D, Lungu C, Baldessarini RJ. Epidemiology of tardive dyskinesia before and during the era of modern antipsychotic drugs. *Handb Clin Neurol.* 2011;100:601-616.
4. Waln O, Jankovic J. An update on tardive dyskinesia: from phenomenology to treatment. *Tremor and other hyperkinetic movements (New York, NY).* 2013;3.
5. Lydiard RB, Gelenberg AJ. Amoxapine--an antidepressant with some neuroleptic properties? A review of its chemistry, animal pharmacology and toxicology, human pharmacology, and clinical efficacy. *Pharmacotherapy.* 1981;1(3):163-178.
6. Carbon M, Hsieh CH, Kane JM, Correll CU. Tardive Dyskinesia Prevalence in the Period of Second-Generation Antipsychotic Use: A Meta-Analysis. *J Clin Psychiatry.* 2017;78(3):e264-e278.
7. Woerner MG, Correll CU, Alvir JM, Greenwald B, Delman H, Kane JM. Incidence of tardive dyskinesia with risperidone or olanzapine in the elderly: results from a 2-year, prospective study in antipsychotic-naive patients. *Neuropsychopharmacology.* 2011;36(8):1738-1746.
8. Tarsy D, Baldessarini RJ. Epidemiology of tardive dyskinesia: is risk declining with modern antipsychotics? *Mov Disord.* 2006;21(5):589-598.
9. Correll CU, Leucht S, Kane JM. Lower risk for tardive dyskinesia associated with second-generation antipsychotics: a systematic review of 1-year studies. *The American journal of psychiatry.* 2004;161(3):414-425.
10. Vijayakumar D, Jankovic J. Drug-Induced Dyskinesia, Part 2: Treatment of Tardive Dyskinesia. *Drugs.* 2016;76(7):779-787.
11. Browne S, Roe M, Lane A, et al. Quality of life in schizophrenia: relationship to sociodemographic factors, symptomatology and tardive dyskinesia. *Acta psychiatrica Scandinavica.* 1996;94(2):118-124.
12. Boumans CE, de Mooij KJ, Koch PA, van 't Hof MA, Zitman FG. Is the social acceptability of psychiatric patients decreased by orofacial dyskinesia? *Schizophrenia bulletin.* 1994;20(2):339-344.
13. Bergman H, Walker DM, Nikolakopoulou A, Soares-Weiser K, Adams CE. Systematic review of interventions for treating or preventing antipsychotic-induced tardive dyskinesia. *Health technology assessment (Winchester, England).* 2017;21(43):1-218.
14. Rana AQ, Chaudry ZM, Blanchet PJ. New and emerging treatments for symptomatic tardive dyskinesia. *Drug Des Devel Ther.* 2013;7:1329-1340.
15. Davis MC, Miller BJ, Kalsi JK, Birkner T, Mathis MV. Efficient Trial Design - FDA Approval of Valbenazine for Tardive Dyskinesia. *The New England journal of medicine.* 2017;376(26):2503-2506.
16. Schooler NR, Kane JM. Research diagnoses for tardive dyskinesia. *Arch Gen Psychiatry.* 1982;39(4):486-487.
17. Miller DD, Eudicone JM, Pikalov A, Kim E. Comparative assessment of the incidence and severity of tardive dyskinesia in patients receiving aripiprazole or haloperidol for the treatment of schizophrenia: a post hoc analysis. *J Clin Psychiatry.* 2007;68(12):1901-1906.

18. Correll CU, Kane JM. One-year incidence rates of tardive dyskinesia in children and adolescents treated with second-generation antipsychotics: a systematic review. *Journal of child and adolescent psychopharmacology.* 2007;17(5):647-656.
19. Woods SW, Morgenstern H, Saksa JR, et al. Incidence of tardive dyskinesia with atypical versus conventional antipsychotic medications: a prospective cohort study. *J Clin Psychiatry.* 2010;71(4):463-474.
20. Maher AR, Theodore G. Summary of the comparative effectiveness review on off-label use of atypical antipsychotics. *Journal of managed care pharmacy : JMCP.* 2012;18(5 Suppl B):S1-20.
21. Patrick M. Neurocrine Biosciences' Valbenazine Launch: What You Need to Know. *Neurocrine Biosciences: The Formula for 2016 Valuations* 2016; http://marketrealist.com/2016/03/neurocrine-biosciences-expected-submit-nda-tardive-dyskinesia-drug-valbenazine-2016/.
22. Zutshi D, Cloud LJ, Factor SA. Tardive Syndromes are Rarely Reversible after Discontinuing Dopamine Receptor Blocking Agents: Experience from a University-based Movement Disorder Clinic. *Tremor and other hyperkinetic movements (New York, NY).* 2014;4:266.
23. Gharabawi GM, Bossie CA, Lasser RA, Turkoz I, Rodriguez S, Chouinard G. Abnormal Involuntary Movement Scale (AIMS) and Extrapyramidal Symptom Rating Scale (ESRS): cross-scale comparison in assessing tardive dyskinesia. *Schizophr Res.* 2005;77(2-3):119-128.
24. Bhidayasiri R, Fahn S, Weiner WJ, Gronseth GS, Sullivan KL, Zesiewicz TA. Evidence-based guideline: treatment of tardive syndromes: report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology.* 2013;81(5):463-469.
25. Lehman AF, Lieberman JA, Dixon LB, et al. Practice guideline for the treatment of patients with schizophrenia, second edition. *The American journal of psychiatry.* 2004;161(2 Suppl):1-56.
26. Vinuela A, Kang UJ. Reversibility of tardive dyskinesia syndrome. *Tremor and other hyperkinetic movements (New York, NY).* 2014;4:282.
27. Glazer WM, Morgenstern H, Schooler N, Berkman CS, Moore DC. Predictors of improvement in tardive dyskinesia following discontinuation of neuroleptic medication. *The British journal of psychiatry : the journal of mental science.* 1990;157:585-592.
28. Glazer WM, Morgenstern H, Doucette JT. The prediction of chronic persistent versus intermittent tardive dyskinesia. A retrospective follow-up study. *The British journal of psychiatry : the journal of mental science.* 1991;158:822-828.
29. Jankelowitz SK. Treatment of neurolept-induced tardive dyskinesia. *Neuropsychiatr Dis Treat.* 2013;9:1371-1380.
30. Erickson JD, Schafer MK, Bonner TI, Eiden LE, Weihe E. Distinct pharmacological properties and distribution in neurons and endocrine cells of two isoforms of the human vesicular monoamine transporter. *Proceedings of the National Academy of Sciences of the United States of America.* 1996;93(10):5166-5171.
31. Muller J, Wissel J, Kemmler G, et al. Craniocervical dystonia questionnaire (CDQ-24): development and validation of a disease-specific quality of life instrument. *Journal of neurology, neurosurgery, and psychiatry.* 2004;75(5):749-753.
32. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med.* 1997;126(5):376-380.
33. Higgins JP, Green S. *Cochrane Collaboration Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011].* The Cochrane Collaboration and John Wiley & Sons Ltd; 2008.

34. 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.
35. Ollendorf DA, Pearson SD. An integrated evidence rating to frame comparative effectiveness assessments for decision makers. *Medical care.* 2010;48(6 Suppl):S145-152.
36. Hauser RA, Factor SA, Marder SR, et al. KINECT 3: A Phase 3 Randomized, Double-Blind, Placebo-Controlled Trial of Valbenazine for Tardive Dyskinesia. *The American journal of psychiatry.* 2017;174(5):476-484.
37. Hauser RA, Factor SA, Marder SR, et al. KINECT 3: A phase 3 randomized, double-blind, placebo-controlled trial of valbenazine for tardive dyskinesia. *American Journal of Psychiatry.* 2017;174(5):476-484.
38. O'Brien CF, Jimenez R, Hauser RA, et al. NBI-98854, a selective monoamine transport inhibitor for the treatment of tardive dyskinesia: A randomized, double-blind, placebo-controlled study. *Movement Disorders.* 2015;30(12):1681-1687.
39. Anderson KE, Stamler D, Davis MD, et al. Deutetrabenazine for treatment of involuntary movements in patients with tardive dyskinesia (AIM-TD): a double-blind, randomised, placebo-controlled, phase 3 trial. *The lancet Psychiatry.* 2017.
40. Fernandez HH, Factor SA, Hauser RA, et al. Randomized controlled trial of deutetrabenazine for tardive dyskinesia the ARM-TD study. *Neurology.* 2017;88(21):2003-2010.
41. Kazamatsuri H, Chien C, Cole JO. Treatment of tardive dyskinesia. I. Clinical efficacy of a dopamine-depleting agent, tetrabenazine. *Archives of general psychiatry.* 1972;27(1):95-99.
42. Grigoriadis D, Comella C, Remington G, Jimenez R, Burke J, Christopher O. Efficacy of valbenazine (NBI-98854) in subjects with tardive dyskinesia: results of a long-term extension study (KINECT 3 extension). *Neuropsychopharmacology Conference: 55th annual meeting of the american college of neuropsychopharmacology, ACNP.* 2016;41:S213-S214.
43. (FDA) USFDA. Ingrezza (valbenazine) Medical Review. 2017; https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/209241Orig1s000MedR.pdf. Accessed August 4, 2017.
44. Bari M, Shiwach R, Jimenez R, Siegert S, O'Brien C. Open-label extension of kinect: A phase 2 study of valbenazine (NBI-98854) for Tardive Dyskinesia. *Neurology.* 2016;86(16).
45. Kane J CC, Wright C, et al. Long-Term Effects of Valbenazine (NBI-98854) in Subjects with Tardive Dyskinesia and Schizophrenia or Schizoaffective Disorder. Paper presented at: College of Psychiatric and Neurologic Pharmacists Annual Meeting2017; Phoenix, AZ.
46. Correll C JR, Liang G, et al. . Efficacy of Valbenazine (NBI-98854) in Treating Subjects with Tardive Dyskinesia and Mood Disorder. Paper presented at: American Psychiatric Association Annual Meeting2017; San Diego, CA.
47. Josiassen R RG, Siegert, S, et al. Valbenazine (NBI-98854) is Effective for Treating Tardive Dyskinesia in Individuals with Schizophrenia or Mood Disorder. Paper presented at: American Psychiatric Association Annual Meeting2016; Atlanta, GA.
48. ClinicalTrials.gov. NBI-98854 for the Treatment of Tardive Dyskinesia in Subjects With Schizophrenia or Schizoaffective Disorder. 2017; https://clinicaltrials.gov/ct2/show/results/NCT01393600?term=NCT01393600&rank=1. Accessed August 25, 2017.
49. Anderson KE SD, Davis MD , et.al. Deutetrabenazine for the Treatment of Tardive Dyskinesia: Results From an Open-Label, Long-Term Study. Paper presented at: The American Psychiatric Association Annual Meeting2017; San Diego, CA.

50. Anderson KE SD, Davis MD , et.al. Long-Term Safety of Deutetrabenazine for the Treatment of Tardive Dyskinesia: Results From an Open-Label, Long-Term Study. Paper presented at: The American Psychiatric Association Annual Meeting2017; San Diego, CA.
51. Jimenez-Shahed J, Fernandez HH, Stamler D, et al. Deutetrabenazine treatment response by concomitant dopamine-receptor antagonists in the phase III, randomized, double-blind, placebo-controlled AIM-TD trial in tardive dyskinesia (TD). *Neurology.* 2017;88(16).
52. Fernandez HH, Hauser RA, Davis MD, et al. Addressing involuntary movements in tardive dyskinesia (AIM-TD): Improvements in clinical global impression of change (CGIC) with deutetrabenazine treatment in moderate to severe tardive dyskinesia (TD). *Neurology.* 2017;88(16).
53. Fernandez HH SD, Hauser RA, et al. Evaluation of Patient-Reported Outcomes in Tardive Dyskinesia Patients With Underlying Psychotic and Mood Disorders in the ARM-TD and AIM-TD Trials. Paper presented at: The American Psychiatric Association Annual Meeting2017; San Diego, CA.
54. Huntington Study G, Frank S, Testa CM, et al. Effect of Deutetrabenazine on Chorea Among Patients With Huntington Disease: A Randomized Clinical Trial. *Jama.* 2016;316(1):40-50.
55. (FDA) USFaDA. Austedo (deutetrabenazine) FDA Medical Review. 2017; https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/208082Orig1s000MedR.pdf. Accessed August 21, 2017.
56. (FDA) USFaDA. Austedo (deutetrabenazine) Prescribing Information. 2017; https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/209885lbl.pdf. Accessed October 27, 2017.
57. Jankovic J, Orman J. Tetrabenazine therapy of dystonia, chorea, tics, and other dyskinesias. *Neurology.* 1988;38(3):391-394.
58. Jankovic J, Beach J. Long-term effects of tetrabenazine in hyperkinetic movement disorders. *Neurology.* 1997;48(2):358-362.
59. Ondo WG, Hanna PA, Jankovic J. Tetrabenazine treatment for tardive dyskinesia: assessment by randomized videotape protocol. *American Journal of Psychiatry.* 1999;156(8):1279-1281.
60. Kenney C, Hunter C, Jankovic J. Long-term tolerability of tetrabenazine in the treatment of hyperkinetic movement disorders. *Movement Disorders.* 2007;22(2):193-197.
61. Paleacu D, Giladi N, Moore O, Stern A, Honigman S, Badarny S. Tetrabenazine Treatment in Movement Disorders. *Clin Neuropharmacol.* 2004;27(5):230-233.
62. Miguel R, Mendonça MD, Barbosa R, et al. Tetrabenazine in treatment of hyperkinetic movement disorders: An observational study. *Therapeutic Advances in Neurological Disorders.* 2017;10(2):81-90.
63. Huntington Study G. Tetrabenazine as antichorea therapy in Huntington disease: a randomized controlled trial. *Neurology.* 2006;66(3):366-372.
64. Frank S. Tetrabenazine as anti-chorea therapy in Huntington disease: an open-label continuation study. Huntington Study Group/TETRA-HD Investigators. *BMC neurology.* 2009;9(62).
65. (FDA) USFaDA. Xenazine (tetrebenazine) package insert. 2008; https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/021894s010lbl.pdf. Accessed August 10, 2017.
66. Chong SA, Remington G, Mahendran R, Chua HC. Awareness of tardive dyskinesia in Asian patients with schizophrenia. *J Clin Psychopharmacol.* 2001;21(2):235-237.

67. Arango C, Adami H, Sherr JD, Thaker GK, Carpenter WT, Jr. Relationship of awareness of dyskinesia in schizophrenia to insight into mental illness. *The American journal of psychiatry.* 1999;156(7):1097-1099.
68. Lambert M, Conus P, Eide P, et al. Impact of present and past antipsychotic side effects on attitude toward typical antipsychotic treatment and adherence. *European psychiatry : the journal of the Association of European Psychiatrists.* 2004;19(7):415-422.
69. Novick D, Haro JM, Suarez D, Perez V, Dittmann RW, Haddad PM. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. *Psychiatry research.* 2010;176(2-3):109-113.
70. Emsley R, Chiliza B, Asmal L, Harvey BH. The nature of relapse in schizophrenia. *BMC psychiatry.* 2013;13:50.
71. Ascher-Svanum H, Faries DE, Zhu B, Ernst FR, Swartz MS, Swanson JW. Medication adherence and long-term functional outcomes in the treatment of schizophrenia in usual care. *J Clin Psychiatry.* 2006;67(3):453-460.
72. Barendregt J. The Half-Cycle Correction: Banish Rather than Explain It. *Medical Decision Making.* 2009;29(4):500-502.
73. Carney CP, Jones L, Woolson RF. Medical comorbidity in women and men with schizophrenia: a population-based controlled study. *Journal of general internal medicine.* 2006;21(11):1133-1137.
74. Shippee ND, Shah ND, Williams MD, Moriarty JP, Frye MA, Ziegenfuss JY. Differences in demographic composition and in work, social, and functional limitations among the populations with unipolar depression and bipolar disorder: results from a nationally representative sample. *Health and quality of life outcomes.* 2011;9:90.
75. Fernandez HH, Stamler D, Davis MD, et al. Improvements in clinical global impression of change with deutetrabenazine treatment in tardive dyskinesia from the ARM-TD and AIM-TD studies. *The American Psychiatric Association 2017 Annual Meeting, San Diego CA; May 20-24, 2017.* 2017.
76. Anderson KE, Stamler D, Davis MD, et al. Long-Term Safety of Deutetrabenazine for the Treatment of Tardive Dyskinesia: Results From an Open-Label, Long-Term Study. *The American Psychiatric Association 2017 Annual Meeting, San Diego CA; May 20-24, 2017.* 2017.
77. Remington G, Comella CL, Grigoriadis D, Liang GS, Burke J, O'Brien CF. Safety and tolerability of valbenazine (NBI-98854) in subjects with tardive dyskinesia: Results of long-term exposure data from three studies. *55th Annual Meeting of the American College of Neuropsychopharmacology, Dec 4-8, 2016; Hollywood, FL.* 2016.
78. Anderson KE, Stamler D, Davis MD, et al. Deutetrabenazine for the treatment of tardive dyskinesia: Results from an open-label, long-term study. *The American Psychiatric Association 2017 Annual Meeting, San Diego CA; May 20-24, 2017.* 2017.
79. Hayes JF, Miles J, Walters K, King M, Osborn DP. A systematic review and meta-analysis of premature mortality in bipolar affective disorder. *Acta psychiatrica Scandinavica.* 2015;131(6):417-425.
80. Saha S, Chant D, McGrath J. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Archives of general psychiatry.* 2007;64(10):1123-1131.
81. Aitken M, Kleinrock M, Pennente K, Lyle J, Nass D, Caskey I. Medicines Use and Spending in the U.S.: A review of 2015 and Outlook to 2020. . *Parsippany, NJ: QuintilesIMS; 2016.* 2016.
82. Physician Fee Schedule Search. 2017. https://www.cms.gov/apps/physician-fee-schedule/search/search-criteria.aspx. Accessed September 12, 2017.

83. Calvert NW, Burch SP, Fu AZ, Reeves P, Thompson TR. The cost-effectiveness of lamotrigine in the maintenance treatment of adults with bipolar I disorder. *Journal of managed care pharmacy : JMCP.* 2006;12(4):322-330.
84. Wang PS, Ganz DA, Benner JS, Glynn RJ, Avorn J. Should clozapine continue to be restricted to third-line status for schizophrenia?: a decision-analytic model. *The journal of mental health policy and economics.* 2004;7(2):77-85.
85. Lenert L, Sturley A, MH R, Chavez S, Morhr P, Rupnow M. Public preferences for health states with schizophrenia and a mapping function to estimate utilities from positive and negative symptom scale scores. *Schizophr Res* 2004;71(1):155-165.
86. Briggs A. Statistical Approaches to Handling Uncertainty in Health Economic Evaluation. *European Journal of Gastroenterology &Hepatology.* 2004;16:551-561.
87. Ascher-Svanum H, Zhu B, Faries D, Peng X, Kinon BJ, Tohen M. Tardive dyskinesia and the 3-year course of schizophrenia: results from a large, prospective, naturalistic study. *The Journal of clinical psychiatry.* 2008;69(10):1580-1588.
88. Rosenheck RA. Evaluating the cost-effectiveness of reduced tardive dyskinesia with second-generation antipsychotics. *The British journal of psychiatry : the journal of mental science.* 2007;191:238-245.
89. Dhir A, Schilling T, Abler V, Potluri R, Carroll B. Estimation of Epidemiology of Tardive Dyskinesia Incidence and Prevalence in the United States. American Psychiatric Association 2017 Annual Meeting; 2017; San Diego, CA.
90. Mihalopoulos C, Harris M, Henry L, Harrigan S, McGorry P. Is early intervention in psychosis cost-effective over the long term? *Schizophrenia bulletin.* 2009;35(5):909-918.
91. (APA) APA. Choosing Wisely: Five Things Physicians and Patients Should Question. 2013; http://www.choosingwisely.org/wp-content/uploads/2015/02/APA-Choosing-Wisely-List.pdf. Accessed September 8, 2017.
92. Society AG. Choosing Wisely: Ten Things Clinicians and Patients Should Question. 2013; http://www.choosingwisely.org/wp-content/uploads/2015/02/AGS-Choosing-Wisely-List.pdf. Accessed September 8, 2017.
93. Anthem Blue Cross Blue Shield. Connecticut Select Drug List. Updated July 2017. https://www11.anthem.com/pharmacyinformation/home.html.
94. Connecticare. Freedom Formulary: 4 Tiers for Plans Purchased on Access Health CT (Connecticut Exchange). Updated August 2017. https://main.myconnecticare.com/~/media/freedomformulary4tier_2017.pdf?la=en.
95. Anthem Blue Cross Blue Shield. Maine Select Drug List. Updated July 2017. https://www11.anthem.com/pharmacyinformation/home.html.
96. Harvard Pilgrim Health Care. 4-Tier Premium Formulary. Last Updated August 2017. https://www.harvardpilgrim.org/hphconnect/formulary/drugLookup.htm?drugName=Tetrabenazine&formularyId=PREMIUM_4_TIER&year=2017&submitBtn=Find.
97. Blue Cross Blue Shield of Massachusetts. 3 Tier Pharmacy Program Look up. *Online.* Last Updated July 2017.
98. Neighborhood Health Plan of Massachusetts. 4 Tier Formulary. 2017. https://fm.formularynavigator.com/MemberPages/pdf/4TFormulary_11417_Brand%20Reference%20Drug_3299.pdf.
99. Tufts Health Plan. Massachusetts Individual and Small Group 3-Tier Drug List. Updated August 2017.

https://fm.formularynavigator.com/MemberPages/pdf/2017MA3TonoffExchange_10796_Full_3051.pdf.
100. Anthem Blue Cross Blue Shield. New Hampshire Select Drug List. Updated July 2017. https://www11.anthem.com/pharmacyinformation/home.html.
101. Blue Cross & Blue Shield of Rhode Island. Individual Markets 5-Tier Prescription Drug List. Updated August 2017. https://www.myprime.com/content/dam/prime/memberportal/forms/AuthorForms/IVL/2017/2017_RI_5T_HealthInsuranceMarketplace.pdf.
102. Neighborhood Health Plan of Rhode Island. NHPRI Rhode Island Health Benefits Exchange 2017 Drug Formulary. Updated August 2017. https://client.formularynavigator.com/Search.aspx?siteCode=8452790730.
103. Blue Cross Blue Shield of Vermont. Formulary Document. 2017. http://www.bcbsvt.com/wps/wcm/connect/ea8c6092-c785-46b9-a118-002f0e1f7a35/bcbsvt-formulary-2017.pdf?MOD=AJPERES.
104. MVP Health Group. Marketplace Formulary: New York- Vermont. Updated June 2017. https://www.mvphealthcare.com/wp-content/uploads/download-manager-files/MVP-Health-Care-Marketplace-Formulary_June-2017.pdf.
105. Agency for Healthcare Research and Quality. *U.S. Preventive Services Task Force Procedure Manual.* 2008.
106. Anderson KE, Stamler D, Davis MD, et al. Addressing involuntary movements in tardive dyskinesia (AIM-TD): Efficacy, safety, and tolerability of fixed-dose deutetrabenazine for the treatment of moderate to severe tardive dyskinesia (TD). *Neurology.* 2017;88(16).
107. Remington G, Comella C, Grigoriadis D, Liang G, Burke J, O'Brien C. Safety and tolerability of valbenazine (NBI-98854) in subjects with tardive dyskinesia: results of long-term exposure data from three studies. *Neuropsychopharmacology Conference: 55th annual meeting of the american college of neuropsychopharmacology, ACNP.* 2016;41:S212-S213.
108. Correll C JR, Le H, et al. Long-Term Effects of Valbenazine (NBI-98854) in Subjects with Tardive Dyskinesia and Mood Disorder. Paper presented at: College of Psychiatric and Neurologic Pharmacists Annual Meeting2017; Phoenix, AZ.
109. Jimenez-Shahed J, Fernandez HH, Stamler D, et al. Deutetrabenazine is associated with an improvement in involuntary movements in patients with tardive dyskinesia (TD) as seen by the high proportion of responders to deutetrabenazine treatment in the AIM-TD study. *Neurology.* 2017;88(16).
110. Anderson KE SD, Davis MD , et.al. Effect of Deutetrabenazine on Quality of Life in Patients With Tardive Dyskinesia in AIM-TD, a 12-Week Double-Blind, Placebo-Controlled Study. Paper presented at: The American Psychiatric Association Annual Meeting2017; San Diego, CA.
111. Sanders GD NP, Basu A. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA.* 2016;316(10).
112. Neurocrine Biosciences I. Data on file. In: Review IfCaE, ed2017.
113. Red Book Online® Search. Truven Health Analytics; 2017. http://www.micromedexsolutions.com.ezp-prod1.hul.harvard.edu/micromedex2/librarian/CS/E7F89E/ND_PR/evidencexpert/ND_P/evidencexpert/DUPLICATIONSHIELDSYNC/A4E796/ND_PG/evidencexpert/ND_B/evidencexpert/ND_AppProduct/evidencexpert/ND_T/evidencexpert/PFActionId/redbook.FindRedBook?navitem=topRedBook&isToolPage=true. Accessed September 1, 2017.

114. Lenert LA, Ziegler J, Lee T, Sommi R, Mahmoud R. Differences in health values among patients, family members, and providers for outcomes in schizophrenia. *Medical care.* 2000;38(10):1011-1021.
115. Peng X, Ascher-Svanum H, Faries D, Conley RR, Schuh KJ. Decline in hospitalization risk and health care cost after initiation of depot antipsychotics in the treatment of schizophrenia. *ClinicoEconomics and outcomes research : CEOR.* 2011;3:9-14.

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.
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		

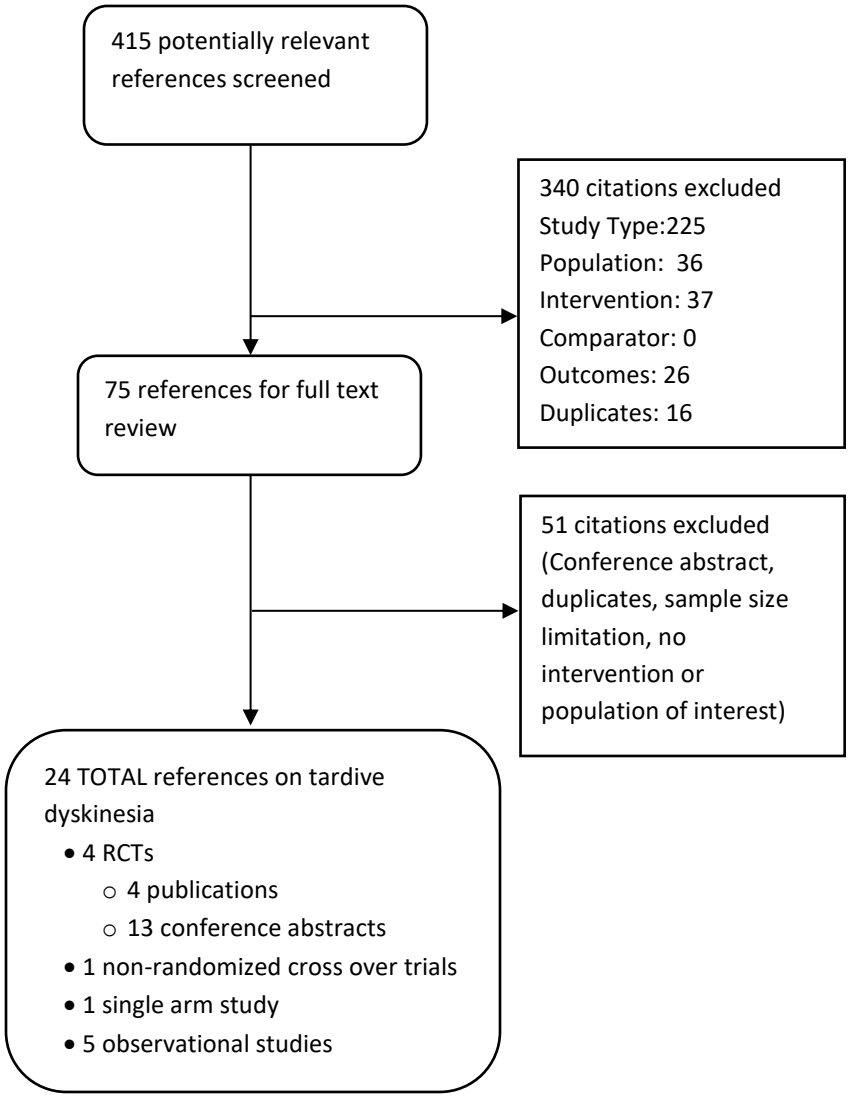
Table A2. Search Strategies of Medline 1996 to Present with Daily Update, Psych INFO and Cochrane Central Register of Controlled Trials

1	**exp tardive dyskinesia/**
2	(tardive adj3 (dyskine$ or diskine$ or syndrome$ or dystonia$)).ti,ab.
3	1 or 2
4	(movement* adj disorder*).mp.
5	((involuntary* or abnormal* or hyperkinetic) adj3 movement*).mp.
6	3 or 4 or 5
7	exp tetrabenazine/
8	(Tetrabenazine or Xenazine).mp.
9	(Deutetrabenazine or Austedo).mp.
10	(Valbenazine or ingrezza).mp.
11	7 or 8 or 9 or 10
12	vesicular monoamine transporter adj3 inhibitor.mp
13	11 or 12
14	6 and 13
15	(animals not (humans and animals)).sh.
16	14 not 15
17	limit 16 to english language
18	(abstract or addresses or autobiography or bibliography or biography or clinical trial, phase I or comment or congresses or consensus development conference or duplicate publication or editorial or guideline or in vitro or interview or lecture 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 videoaudio media).pt.
19	17 not 18
20	remove duplicates from 19

Table A3. Embase Search Strategy

#1	**'tardive dyskinesia'/exp**
#2	(tardive NEAR/3 (dyskine* OR diskine* OR dystonia* OR syndrome*)):ab,ti
#3	**#1** OR **#2**
#4	'movement disorder*':ab,ti
#5	((involuntary* OR abnormal*) NEAR/3 movement*):ab,ti
#6	#3 OR #4 OR #5
#7	'tetrabenazine'/exp OR tetrabenazine:ab,ti OR xenazine:ab,ti
#8	'deutetrabenazine'/exp OR deutetrabenazine:ab,ti OR austedo:ab,ti
#9	'valbenazine'/exp OR valbenazine:ab,ti OR ingrezza:ab,ti
#10	#7 OR #8 OR #9
#11	'vesicular monoamine transporter' NEAR/3 inhibitor*
#12	#10 OR #11
#13	#6 AND #12
#14	'animal'/exp OR 'nonhuman'/exp OR 'animal experiment'/exp
#15	'human'/exp
#16	#14 AND #15
#17	#14 NOT #16
#18	#13 NOT #17
#19	#18 AND ('chapter'/it OR 'conference review'/it OR 'editorial'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)
#20	#18 NOT #19
#21	#20 AND [english]/lim
#22	#21 AND [medline]/lim
#23	#21 NOT #22

Figure A1. PRISMA Flow Chart Showing Results of Literature Search for Tardive Dyskinesia



Appendix B. Clinical Guidelines

Table B1. American Academy of Neurology Guidelines: Level of Evidence for TD Interventions[24]

Insufficient	Weak	Moderate
Altering DRBA Therapy		
Withdrawing DRBA		
Switching from older (first-generation) to newer (second-generation or atypical) anti-psychotic DRBA		
Pharmacologic Therapies		
First-generation anti-psychotics Although haloperidol and thiopropazate possibly reduce TDS, they are not recommended because of the competing risk of akinetic-rigid syndrome.		
Second-generation anti-psychotics Because neuroleptic agents may themselves cause TDS and may mask its symptoms rather than treat it, these drugs cannot be recommended for TDS treatment		
Dopamine Depleting Agents		
Reserpine or α-methyldopa	Tetrabenazine Possibly reduces symptoms and may be considered. No evidence that long-term TBZ administration induces TDS, but it can cause parkinsonism	
Cholinergic and Anticholinergic Drugs		
All other cholinergic and anticholinergic drugs	Galantamine is possibly ineffective and should not be considered	
Antioxidants		
Vitamin E, melatonin, vitamin B6, selegiline, and yi-gan san	Eicosapentaenoic acid possibly ineffective and should not be considered	Ginkgo biloba is likely helpful for patients, but data is limited to those with schizophrenia
Gaba Agents		
Baclofen		Clonazepam should be considered for short term treatment
Calcium Channel Blockers		

Nifedipine		Diltiazem should not be considered a treatment
Other Pharmacologic Agents or Health System Interventions		
Acetazolamide	Amantadine may be used with neuroleptics for short term use	
Buspirone		
Biperiden (Akineton) Discontinuation		
Chemodenervation with botulinum toxin		
Dopamine agonists: bromocriptine		
Electroconvulsive therapy		
Levetiracetam		
Pallidal deep brain stimulation		

Appendix C. Public and Representative Private Insurer Coverage Policies

Table C1. Coverage Policies for New England Commercial Payers

	Connecticut		Maine		Massachusetts			New Hampshire		Rhode Island		Vermont	
	Anthem (Wellpoint Inc Group)[93]	Connecticare[94]	Anthem (Wellpoint Inc Group)[95]	HPHC Maine[96]	BCBS of MA[97]	Neighbor-hood Health Plan[98]	Tufts Health Plan[99]	Anthem (Wellpoint Inc Group)[100]	HPHC New Hampshire[96]	BCBS of RI[101]	Neighbor-hood Health Plan of RI[102]	BCBS of VT[103]	MVP Grp[104]
Tetrabenazine													
Covered	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tier	No	4	2	2	1	No	1	2	1	5	4	1	2
PA	No	Yes	Yes	No	No	No	Yes	Yes	No	Yes	No	No	No
Notes:		SP*		SP	SP		SP		SP		Branded preferred		SP
Valbenazine													
Covered	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No
Tier	No	4	No	4	3	No	No	No	No	No	No	No	No
PA	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No
Deutetrabenazine													
Covered	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No
Tier	No	4	No	4	3	No	No	No	No	No	No	No	No
PA	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No
Other Tardive Dyskinesia Off-Indication Coverage													
Amantadine	Tier 1		Tier 2	Tier 2	Tier 1			Tier 2	Tier 1				
Clonazepam	Tier 1		Tier 1	Tier 2	Tier 1			Tier 1	Tier 1				

Appendix D. Previous Systematic Reviews and Technology Assessments

Previous Systematic Reviews

We identified one systematic review on tetrabenazine for use in a variety of hyperkinetic movement disorders including tardive dyskinesia, and one systematic review on valbenazine for tardive dyskinesia. An additional systematic review assessing the impact of neuropletic reduction/cessation and neuroleptic medication use as specific treatments for TD was also identifed. These reviews are summarized below.

Chen JJ, Ondo WG, Dashtipour K, Swope DM. Tetrabenazine for the Treatment of Hyperkinetic Movement Disorders: A Review of the Literature. Clinical Therapeutics. 2012; (34):1487-1504

In this review, Chen and colleagues describe the clinical efficacy and tolerability of tetrabenazine in the management of various hyperkinetic movement disorders including dystonia, Huntington's chorea, tardive dyskinesia, and tic disorders. The researchers identified nine retrospective studies and two prospective studies of tetrabenazine in patients with tardive dyskinesia between 1974 and 2008. Only two studies reported statistical values, however in nine of the 11 studies, tetrabenazine demonstrated a clinical benefit for tardive dyskinesia symptoms. The researchers noted that additional randomized, placebo-controlled studies are necessary to demonstrate the efficacy of tetrabenazine for TD. Adverse events from tetrabenazine was found to be dose dependent and age related. Some common adverse events identified in the trials include somnolence, insomnia, akathisia, depression, parkinsonism, and fatigue. Tetrabenazine also includes a black box warning regarding depression and suicidality in its FDA approved indication for Huntington's chorea.

Citrome L. Valbenazine for tardive dyskinesia: A systematic review of the efficacy and safety profile for this newly approved novel medication—What is the number needed to treat, number needed to harm and likelihood to be helped or harmed? International Journal of Clinical Practice. 2017; 3-14

In this review, Citrome describes the efficacy, safety, and tolerability of valbenazine in the treatment of tardive dyskinesia. A total of 13 trials for valbenazine were identified in which seven trials were conducted in patients with TD and included in this review. An in-depth review of two of the trials (KINECT 2 & 3 trials), including additional analyses of the number needed to treat (NNT) and number needed to harm (NNH) was conducted. Valbenazine was shown to result in a statistically significant greater AIMS response, as defined by ≥50% reduction in AIMS score from baseline, compared to placebo in both trials. The resulting pooled NNT was 5. Additionally, a pooled NNT of 5 was estimated based on responders from the clinical global impression of change

(CGIC) of tardive dyskinesia scale. Discontinuation rates because of an AE from the pooled study was 2.9% for patients treated with valbenazine versus 1.6% for placebo-treated patients, resulting in a non-significant NNH of 76. Although, the efficacy and tolerability of VBZ were established in these randomized, placebo-controlled clinical trials, the author noted that head-to-head comparisons with other VMAT2 inhibitors would be needed to fully evaluate the role of valbenazine in the treatment of TD.

Soares-Weiser K, Rathbone J. Neuroleptic reduction and/or cessation and neuroleptics as specific treatments for tardive dyskinesia (Review). Cochrane Database of Systematic Reviews. 2006; (1):1-39

In this review, Soares-Weiser and colleagues sought to assess the impact of neuroleptic reduction or cessation in reducing TD symptoms. Two randomized, double-blind, controlled trials comparing neuroleptic reduction to continuing neuroleptic medications were identified. Both studies found no association between neuroleptic reduction and improvement in TD symptoms. No randomized controlled trials relevant to neuroleptic cessation as a treatment for TD was identified.

Appendix E. Ongoing Studies

Title, Trial Sponsor, ClinicalTrials.gov Identifier	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
Deutetrabenazine					
Reducing Involuntary Movements in Tardive Dyskinesia (RIM-TD) Auspex Pharmaceuticals, Inc. NCT02198794	Phase III Long-Term Safety Study Open-Label Estimated Enrollment: 343	1. Deutetrabenazine tablets dose titrated for 6 weeks until optimal dose is reached	Inclusion Criteria • History of using a dopamine receptor blocking agent for at least 3 months. • History of being compliant with prescribed medications. • Subjects with underlying psychiatric diagnosis are stable and have no change in psychoactive medications. • Be in good general health and is expected to attend all study visits and complete study assessments. Exclusion Criteria • Serious untreated or undertreated psychiatric illness. • Evidence of hepatic or renal impairment. • History of alcohol or substance abuse in the previous 12 months. • Unstable or serious medical illness. • Neurological condition other than tardive dyskinesia.	Primary Outcome Measures • Safety [Time Frame: 159 weeks] : Incidence of AEs, serious AEs, severe AEs, drug related AEs, AEs leading to withdrawal. Secondary Outcome Measures [Time Frame: 158 weeks] • Change in AIMS score from Baseline • Quality of Life Exploratory Efficacy Assessment • PGIC Exploratory Efficacy – treatment success based on PGIC • AIMS Exploratory Efficacy Assessment - % change in AIMS score • AIMS responders Exploratory Efficacy Assessment - Proportion of responders based on AIMS change from baseline	September 2019

Title, Trial Sponsor, ClinicalTrials.gov Identifier	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
Tetrabenazine					
Xenazine in Late Dyskinetic Syndrome With Neuroleptics (Xeladys) Centre Hospitalier Universitaire, Amiens NCT01543321	Phase III RCT Parallel Group Double-blind Estimated Enrollment:54	1.Tetrabenazine • 5-week titration to a maximum dose of 200 mg / day • 5 weeks at stable dose • 2 weeks in wash-out 2. Placebo • 5-week titration to a maximum dose of 200 mg / day • 5 weeks at stable dose • 2 weeks in wash-out	Inclusion Criteria • QTc < 450 ms for men and <470 for women • MADRS < 18 • Adult (age over 18) or adult under judicial protection (tutor or curator). • Patient with late dyskinetic syndrome with neuroleptics yielding functional disability and/or impact in every day life, according to the investigator, and/or the patient and/or the patient's family. Exclusion Criteria • Insanity according to the DSM IV and MMS < 24 • Predominant akathisia • Renal failure • Congenital galactosemia, glucose malabsorption or lactase deficiency • Drugs: Non-selective MAOIs, dopaminergic (or other antiparkinsonian)	Primary Outcome Measures • Changes in Extrapyramidal Symptoms Rating Scale (ESRS) [Time Frame: 10 weeks after randomization] Secondary Outcome Measures • AIMS improvement • Changes in Quality of life • Changes in intermediate ESRS and post-treatment ESRS • Tolerance • CGI amelioration • Changes in Sub-score ESRS-II	June 2017

Title, Trial Sponsor, ClinicalTrials.gov Identifier	Study Design	Treatment Arms	Patient Population	Key Outcomes	Estimated Completion Date
Valbenazine					
Rollover Study for Continuing Valbenazine (NBI-98854) Administration for the Treatment of Tardive Dyskinesia Neurocrine Biosciences NCT02736955	Phase III Interventional Open label Estimated Enrollment: 150	1. Fixed dose of valbenazine administered once daily for up to 72 weeks	Inclusion Criteria • Have participated in and completed the Kinect 3 or Kinect 4 Phase 3 study. • Have a negative urine drug screen for amphetamines, barbiturates, benzodiazepine, phencyclidine, cocaine, opiates, or cannabinoids. • If using maintenance medication(s) for schizophrenia or schizoaffective disorder, mood disorder, or other conditions, be on stable doses. Exclusion Criteria • Have a known history of long QT syndrome or cardiac arrhythmia. • Have a known history of substance dependence, substance (drug) or alcohol abuse. • Have a known history of neuroleptic malignant syndrome. • Have a blood loss ≥550 mL or donated blood within 30 days prior to Baseline.	Primary Outcome Measures • Safety and tolerability measurements (incidence and types of adverse events, physical examination, laboratory tests) [Time Frame: Up to 72 weeks] Secondary Outcome Measures • CGI-TD-Severity • Patient Satisfaction Questionnaire • Social Functioning Scale (SFS)	June 2017

Source: www.ClinicalTrials.gov (NOTE: studies listed on site include both clinical trials and observational studies)

Appendix F. Comparative Clinical Effectiveness Supplemental Information

We performed screening at both the abstract and full-text level. Three investigators screened all abstracts identified through electronic searches according to the inclusion and exclusion criteria described in the published protocol (https://osf.io/q6hxt/). 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 the agents of interest. These included manufacturer submissions 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 F2)[105] 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, intent-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. Intent-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, intent-to-treat analysis is lacking.*

Data Extraction

Three reviewers extracted key information from the full set of accepted studies. Extracted data were reviewed for logic, and a random proportion of data were validated by a third investigator for additional quality assurance. Summary tables of extracted data are available in Appendix G.

Table F1. Abnormal Involuntary Movement Scale (AIMS)

MOVEMENT RATINGS: Rate highest severity observed. Rate movements that occur upon activation one less than those observed spontaneously. Circle movement as well as code number that applies.		RATER Date	RATER Date	RATER Date	RATER Date
Facial and Oral Movements	**1. Muscles of Facial Expression** e.g. movements of forehead, eyebrows, periorbital area, cheeks, including frowning, blinking, smiling, grimacing	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
	2. Lips and Perioral Area e.g., puckering, pouting, smacking	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
	3. Jaw e.g. biting, clenching, chewing, mouth opening, lateral movement	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
	4. Tongue Rate only increases in movement both in and out of mouth. NOT inability to sustain movement. Darting in and out of mouth.	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Extremity Movements	**5. Upper (arms, wrists, hands, fingers)** Include choreic movements (i.e., rapid, objectively purposeless, irregular, spontaneous) athetoid movements (i.e., slow, irregular, complex, serpentine). DO NOT INCLUDE TREMOR (i.e., repetitive, regular, rhythmic)	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
	6. Lower (legs, knees, ankles, toes) e.g., lateral knee movement, foot tapping, heel dropping, foot squirming, inversion and eversion of foot.	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Trunk Movements	**7. Neck, shoulders, hips** e.g., rocking, twisting, squirming, pelvic gyrations	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
0 = None, normal; 1 = Minimal, may be extreme normal; 2 = Mild; 3 = Moderate; 4 = Severe *Note: Although a typical AIMS contains up to 14 items, the valbenazine and detetrabenazine trials only used items 1-7 to rate AIMS scores, but includes all potentially affected body parts*					

Table F2: Valbenazine: Subgroup Analysis of AIMS Change by Psychiatric Diagnosis

Trials	AIMS Reduction From Baseline	
	Schizophrenia/Schizoaffective Subgroup	Mood Disorder Subgroup
KINECT 3 Trial[45,46]		
Valbenazine 80mg/D	-2.9†	-3.6†
Valbenazine 40mg/D	-1.6†	-2.4
Placebo	+0.3	-0.7
KINECT 2 Trial[47]		
Valbenazine (Max 75mg/D)	-3†	-4.5†
Placebo	-1	-1.4

†p value≤0.05 vs. placebo

Table F3. Changes in MADRS and YMRS from Baseline at Weeks Six and 48 (Mood Disorder Subgroup)

Mood Disorder			
Psychiatric Scale	80mg VBZ	40mg VBZ	PBO
6 Week Data			
MADRS, Mean change	-1.5	0.0	1.2
YMRS, Mean change	-1.4	-0.4	0.5
48 Week Data			
MADRS, Mean change (SEM)	0.6 (-0.6, 1.9)	-0.2 (-1.1, 0.7)	NA
YMRS, Mean change (SEM)	-1.4 (-1.9, -0.7)	-1.2 (-1.8, -0.6)	NA

Negative sign indicates fewer psychiatric symptoms reported. MADRS - Montgomery-Åsberg Depression Rating Scale, YMRS - Young Mania Rating Scale; NA – Not applicable.

Table F4. Changes in PANSS and CDSS from Baseline at Weeks Six and 48 (Schizophrenia/Schizoaffective Disorder Subgroup)

Schizophrenia/Schizoaffective Disorder			
Psychiatric Scale	80mg Vbz	40mg Vbz	Pbo
6 Week Data			
Panss Positive, Mean Change	-0.3	-0.5	-0.0
Panss Negative, Mean Change	0.5	-0.0	-0.0
Panss General Psychopathology, Mean Change	-0.8	-1.3	-0.2
Cdss, Mean Change	-0.4	-0.5	-0.1
48 Week Data			
Panss Positive, Mean Change (Sem)	-1.7 (-2.2, -1.2)	-0.6 (-1.3, 0.1)	Na
Panss Negative, Mean Change (Sem)	-1.2 (-1.7, -0.7)	0.7 (-0.1, 1.4)	Na
Panss General Psychopathology, Mean Change (Sem)	-3.7 (-4.5, -2.9)	-0.5 (-1.7, 0.6)	Na
Cdss, Mean Change (Sem)	-0.3 (-0.9, -0.1)	-0.5 (-0.7, 0.0)	Na

Negative Sign Indicates Fewer Psychiatric Symptoms Reported. Panss - Positive And Negative Syndrome Scale; Cdss - Calgary Depression Scale For Schizophrenia; Na – Not Applicable

Table F5: Deutetrabenazine: Subgroup Analysis of AIMS Change by Baseline DRBA Use

Trials	AIMS Reduction from Baseline	
	Not receiving DRBA	**Receiving DRBA**
AIM-TD[39]		
DTBZ 12mg/day	-2.4*	-2
DTBZ 24mg/day	-3.1*	-3.2*
DTBZ 36mg/day	-3.1*	-3.4*
Placebo	0	-1.7

†p value≤0.001 vs. placebo; *p value≤0.05

Table F6: Deutetrabenazine: Subgroup Analysis of CGIC Treatment Success by Baseline DRBA

Trials	CGIC Treatment Success	
	Not Receiving DRBA	**Receiving DRBA**
AIM-TD[39]		
DTBZ 12mg/day	27	29
DTBZ 24mg/day	58†	46
DTBZ 36mg/day	60†	34
Placebo	8	31

†p value≤0.05 vs. placebo

Table F7. Deutetrabenazine: Mcdq-24 Across Trials at 12 Weeks

Trials	LS Mean Change In Mcdq-24 Score	LS Mean Difference Vs Placebo
AIM-TD[39]		
DTBZ 12mg/day	-5.9	1.2
DTBZ 24mg/day	-10.6	-3.5
DTBZ 36mg/day	-11.6	-4.4
Placebo	-7.1	--
ARM-TD[106]		
DTBZ	-11.1	-2.8
Placebo	-8.3	--

All active treatements not significant compared to placebo

Table F8: Tetrabenazine: Summary of Evidence

Study	Design/Treatment	Primary Outcome Measure	Findings	Study Procedures
Kazamatsuri 1972	Non-randomized cross over study PBO for 4 weeks, TBZ for 6 weeks, and PBO for another 2 weeks (N=24)	Mean oral dyskinesia frequency/minute	Vs. placebo Mean difference TBZ (after 6 weeks): 19.2 (p<0.0005 vs. PBO)	All neuroleptics were withdrawn 4 weeks prior to TBZ initiation
Ondo 1999	Open label single blind TBZ starting dose of 25mg/d to maximum of 100mg/d (N=20)	Blinded AIMS change from baseline	TBZ: Mean change after 20 weeks: -9.7 (p<0.001)	Patients stopped taking offending medications for at least 30 days before study initiation
Jankovic 1988	Retrospective (18 months†) TBZ starting dose of 25mg/d to maximum of 100mg/d (N=44)	Global Response Scale: composite score by patients, caretakers & investigators. Ranges from 1 (marked improvement) to 5 (worsening)	TBZ: Average score: 2.3; 93% had marked or moderate improvement	Many patients were also being treated concurrently with other medications
Jankovic 1997	Retrospective (28.9 months†) TBZ starting dose of 25mg/d to maximum of 100mg/d (N=93)	Global Response scale (see Jankovic 1988 above)	TBZ: 93% had marked or moderate improvement	Many patients were also being treated concurrently with other medications
Paleacu 2004	Retrospective TBZ starting dose of 25mg/d to maximum of 150mg/d (N=17)	Clinical Global Impression of Change (1 to 7): composite score developed with patients	TBZ: 41% had marked or moderate improvement	Chart review
Kenney 2007	Retrospective (28 months†) TBZ (dose range 12.5mg/d-300mg/d) (N=149)	Global Response scale (see Jankovic 1988 above)	TBZ: 84% had marked or moderate improvement	Chart review
Miguel 2017	Retrospective (49 months†) TBZ starting dose of 12.5mg/d to maximum of 225mg/d (N=35)	Subjective Clinical improvement	TBZ: 77% were classified as responders	Chart review

†Mean duration of follow up

Table F9. Tetrabenazine: Adverse Events at 12 weeks and 80 weeks

	Any AE≥1 (%)	Saes≥1 (%)	Discontinuation Due To AE (%)	Somnolence (%)	Insomnia (%)	Depression (%)	Suicidal Ideation (%)
12-Week Placebo Controlled Studies Of HC[63]							
TBZ (25mg-100mg/D)	91	7.4	9.3	31.5	25.9	14.8	NR (1 Suicide)
Placebo	70	0	3.3	3.3	0	0	NR
80-Week Long Term Extension Study[64]							
TBZ (Max 200mg/D)	74.6	16	4	24	13	22.6	NR (1 Suicide Attempt)

Table F10. Tetrabenazine: Adverse Events from Retrospective Studies

Side Effects	Jankovic 1988	Jankovic 1997	Kenney 2007	Shen 2013	Miguel 2017
Number Of Patients	217	400	448	145	108
Dose/Day	25mg-100mg	25mg-100mg	12.5mg-300mg	12.5mg-300mg	12.5mg-225mg
Mean Follow Up	18 Months	29 Months	28 Months	37 Months	49 Months
Somnolence/Fatigue	13%	37%	25%	45%	5%
Parkinsonism	24%	29%	15%	14%	52%
Depression	11%	15%	8%	27%	5%
Insomnia	5%	11%	5%	28%	NR
Nervousness/Anxiety	7%	10%	5%	21%	4%
Akathisia	5%	10%	8%	12%	2%
Nausea/Vomiting	2.3%	5%	6%	NR	1%

Appendix G. Evidence Tables

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Valbenazine (VBZ)						
O'Brien *Mov. Disord.* 2015 [38] **KINECT 2** **Fair**	Phase II, RCT, double-blind, parallel-group, placebo-controlled dose-titration study 22 sites across USA 6 weeks	*mITT population* 1) Placebo, n=44 2) VBZ, n=45 once daily starting dose of 25mg increased in increments of 25mg/2weeks to a maximum of 75mg	Inclusions: Medically and psychiatrically stable male & female subjects 18-85 years old with moderate to severe TD and a clinical diagnosis of schizophrenia, schizoaffective disorder, mood disorder or gastrointestinal disorder	Mean age (SD) 1) 55.6 (9.8) 2) 56.7 (10.8) Mean age at TD diagnosis, (SD) 1) 49.5 (12.1) 2) 48.9 (13.0) Male/Female, n (%) 1) 27 (55.1)/ 22 (44.9) 2) 30 (58.8)/ 21 (41.2) AIMS at baseline, mean(SD) 1) 7.9 (4.5) 2) 8.0 (3.5) Schizophrenia or schizoaffective disorder/mood disorder/gastrointestinal disorder, % 1) 61.2/36.7/2.0 2) 54.9/39.2/5.9	*Week 6* Mean change in AIMS (SD) 1) -1.1 (3.7) 2) -3.6 (3.5) Least Squares Mean Change in AIMS from baseline (SEM) 1) -0.2 (1.1) 2) -2.6 (1.2) CGI-TD response, n (%) 1) 7 (15.9) 2) 30 (66.7) PGIC response, n (%) 1) 14 (31.8) 2) 26 (57.8)	*Week 6* Overall treatment-emergent adverse events (TEAEs), % 1) 32.7% 2) 49.0% Fatigue, n (%) 1) 2 (4.1) 2) 5 (9.8) Headache, n (%) 1) 2 (4.1) 2) 5 (9.8) Somnolence, n (%) 1) 1 (2.0) 2) 3 (5.9)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Hauser, *Am J Psychiatry*, 2017 [37] **KINECT 3** **Good**	Phase III, RCT, double-blind, placebo-controlled trial 59 centers in United States 2 centers in Canada 2 centers in Puerto Rico 6 weeks	*ITT population* = received ≥1 dose of study drug & had ≥1 post-baseline AIMS assessment 1) Placebo, n=76 2) VBZ 40 mg/day, n=70 3) VBZ 80mg/day, n=79	Inclusion: Participants 18-85 years old with medically stable diagnosis of schizophrenia, schizoaffective disorder, or mood disorder for ≥ 3 months prior to screening; had moderate or severe DRA-induced TD for ≥ 3 months prior to screening	Mean age (SD) 1) 57.0 (10.5) 2) 55.3 (8.5) 3) 56.0 (10.1) Baseline AIMS, mean (SD) 1) 9.9 (4.3) 2) 9.7 (4.1) 3) 10.4 (3.6) Male, n (%) 1) 42 (55.3) 2) 42 (58.3) 3) 39 (49.4) White/Black, n (%) 1) 43 (56.6) /29 (38.2) 2) 41 (56.9) /26 (36.1) 3) 44 (55.7) /32 (40.5) Schizophrenia & schizoaffective disorder/mood disorder, n (%) 1) 50 (65.8)/26 (34.2) 2) 48 (66.7)/24 (33.3) 3) 52 (65.8)/27 (34.2) Concomitant medications Any Antipsychotic, n (%) 1) 62 (82.9) 2) 66 (91.7) 3) 65 (82.3)	*Week 6* Least Squares Mean Change in AIMS from baseline 1) -0.1 (NS) 2) -1.9, p=0. 0021 3) -3.2, p<0.001 ≥50% AIMS response reduction, % 1) 8.7 (NS) 2) 23.8, p=0.02 3) 40.0, p<0.001 CGI-TD Score 1) 3.2 (NS) 2) 2.9, p=0.074 3) 2.9, p=0.056	*Week 6* Any AE, n (%) 1) 33 (43.4) 2) 29 (40.3) 3) 40 (50.6) Discontinuation due to AEs, n (%) 1) 4 (5.3) 2) 4 (5.6) 3) 5 (6.3) Death, n (%) 1) 0 (0) 2) 0 (0) 3) 1 (1.3) Somnolence, n (%) 1) 3 (3.9) 2) 4 (5.6) 3) 4 (5.1) Akathisia, n (%) 1) 1 (1.3) 2) 3 (4.2) 3) 2 (2.5) Suicidal Ideation, n (%) 1) 4 (5.3) 2) 3 (4.2) 3) 1 (1.3)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Grigoriadis, *ACNP, 2016* [42] **(KINECT 3 Extension)** **POSTER ABSTRACT** **Good**	Phase III, RCT, double-blind, placebo-controlled (DBPC) followed by a long-term extension (LTE), and Drug-free washout periods DBPC: 6 weeks LTE: 42 weeks followed by 4 weeks washout period	*DBPC Period* 1) PBO, n=76 2) VBZ 40 mg/day, n=70 3) VBZ 80mg/day, n=79 *LTE and Washout periods* 2) VBZ 40 mg/day, n=97 (combined w/ prior placebo group) 3) VBZ 80mg/day n=101 (combined w/ prior placebo group)	*See Hauser, Am J Psychiatry, 2017* [36]	*See Hauser, Am J Psychiatry, 2017* [36]	*Week 48* AIMS score mean change 2) -2.94‡ 3) -4.81‡ CGI-TD mean score 2) 2.41‡ 3) 2.09‡ ≥50% AIMS response reduction, % 2) 28.3 3) 52.4 CGI-TD response, (%) 2) 59.0 3) 76.2 ‡approximate values estimated from curves with digitization software	NR See *Correll, CPNP, 2017* [46]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Remington, *ACNP*, 2016 [107] **POSTER ABSTRACT** **Good**	Pooled LTE population included VBZ-treated subjects from 3 studies KINECT: 50mg/day, 6-week double blind, placebo-controlled (DBPC) period, 6 week open-label treatment period; KINECT 3: 80 or 40mg/day, 6-week DBPC, 42-week double-blind extension period; KINECT 4: 80 or 40mg/day, 48-week open-label treatment Mean duration of exposure: 6 months	1) LTE VBZ 40mg/day, n=197 2) LTE VBZ 80mg/day, n=230	Inclusions: DSM-IV diagnosis of moderate to severe DRA-induced TD for ≥3 months prior to screening; DSM-IV diagnosis of schizophrenia, schizoaffective disorder, or mood disorder and Brief Psychiatric Rating Scale < 50 at screening with stable psychiatric status and stable doses of concomitant medications	*Safety Population* Mean Age (%<65years) 1) 56.2 (81.7) 2) 56.9 (81.3) Male, % 1) 59.4 2) 52.2 White/Black, % 1) 56.3/37.6 2) 65.7/31.3 Mean Age at TD diagnosis 1) 48.6 2) 48.2 Current schizophrenia & schizoaffective disorder/ mood disorder, % 1) 76.6/23.4 2) 67.4/32.6 Any Concomitant antipsychotic, % 1) 88.3 2) 83.0	NR	*Safety Population* Any treatment emergent AE, % 1) 61.0 2) 71.3 All subject Any serious AE, % 1) 11.5 2) 16.5 Discontinuation due to AE, % 1) 16.0 2) 13.5 AE leading to dose reduction, % 1) 5.0 2) 8.3 Somnolence, % 1) 7.5 2) 5.2

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Correll, *CPNP*, 2017 [108] **KINECT 3 Extension** **POSTER ABSTRACT** ***See Grigoriadis, ACNP, 2016*** [42]	*See Grigoriadis, ACNP, 2016* [42]	*DBPC Period: Mood Disorder Subgroup* 1) PBO, n=26 2) VBZ 40 mg/day, n=24 3) VBZ 80mg/day, n=27 *LTE and Washout periods: Overall population* 2) VBZ 40 mg/day, n=97 (combined w/ prior placebo group) 3) VBZ 80mg/day n=101 (combined w/ prior placebo group)	*See Grigoriadis, ACNP, 2016* [42]	*Mood Disorder Subgroup* Mean Age (SD) 1) 57.4 (11.6) 2) 54.7 (9.1) 3) 54.5 (11.1) Mean Age at TD diagnosis, mean (SD) 1) 51.6 (11.8) 2) 48.7 (9.3) 3) 47.6 (10.4) Male, n (%) 1) 8 (30.8) 2) 11 (45.8) 3) 10 (37.0) White, n (%) 1) 23 (88.5) 2) 18 (75.0) 3) 17 (63.0) AIMS score, mean (SD) 1) 11.2 (3.6) 2) 11.4 (3.5) 3) 10.9 (3.8)	*Week 48* AIMS score mean change 2) -4.2 3) -5.8 CGI-TD mean score 2) 2.2 3) 2.0 ≥50% AIMS response reduction, % 2) 33.3 3) 56.0 CGI-TD response, n (%) 2) 61.0 3) 80.0	*LTE period, Overall Safety Population (≥1 dose of assigned drug)* Any treatment emergent AE, n (%) 2) 60 (61.9) 3) 77 (76.2) Discontinuation due to AE, n (%) 2) 13 (13.4) 3) 18 (17.8) Somnolence, n (%) 2) 3 (3.1) 3) 4 (4.0)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Kane, *CPNP*, 2017 [45] **KINECT 3 Extension** **POSTER ABSTRACT** ***See Grigoriadis, ACNP, 2016***[42]	*See Grigoriadis, ACNP, 2016*[42]	*Schizophrenia/Schizoaffective Disorder subgroup* *DBPC Period* 1) PBO, n=50 2) VBZ 40 mg/day, n=48 3) VBZ 80mg/day, n=52 *VE and Washout periods* 2) VBZ 40 mg/day, n=97 (combined w/ prior placebo group) 3) VBZ 80mg/day n=101 (combined w/ prior placebo group)	*See Grigoriadis, ACNP, 2016*[42]	*Schizophrenia/Schizoaffective Disorder subgroup* Mean age (SD) 1) 56.8 (10.0) 2) 55.6 (8.3) 3) 56.8 (9.5) Mean Age at TD diagnosis (SD) 1) 47.7 (9.6) 2) 47.2 (11.2) 3) 47.6 (13.6) Male, n (%) 1) 34 (68.0) 2) 31 (64.6) 3) 29 (55.8) White/Black, n (%) 1) 20 (40.0)/ 27 (54.0) 2) 23 (47.9)/ 21 (43.8) 3) 27 (51.9)/ 24 (46.2) AIMS score, mean (SD) 1) 9.3 (4.5) 2) 8.8 (4.2) 3) 10.1 (3.5)	*Week 48: Schizophrenia/Schizoaffective Disorder subgroup* AIMS score mean change 2) -2.5 3) -4.2 CGI-TD mean score 2) 2.4 3) 2.2	*Overall safety population* *See Correll, CPNP, 2017* [46]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Josiassen, *APA*, 2016 [47] **KINECT 2** **POSTER ABSTRACT** ***See O'Brien Mov. Disord. 2015***[38]	*See O'Brien Mov. Disord. 2015*[38]	*Schizophrenia & schizoaffective subgroup* 1) PBO, n=27 2) VBZ, n=26 once daily starting dose of 25mg increased in increments of 25mg/2weeks to a maximum of 75mg *Mood disorder subgroup* 1) Placebo, n=16 2) VBZ, n=19 once daily starting dose of 25mg increased in increments of 25mg/2weeks to a maximum of 75mg	*See O'Brien Mov. Disord. 2015*[38]	*See O'Brien Mov. Disord. 2015*[38]	*Week 6* *Schizophrenia & schizoaffective subgroup* AIMS mean score change 1) -1.0 2) -3.0, p<0.05 CGI-TD mean score 1) 3.1 2) 2.4 CGI-TD responders, % 1) 14.8 2) 61.5 *Mood disorder subgroup* AIMS mean score change 1) -1.4 (NS) 2) -4.5, p<0.05 CGI-TD mean score 1) 3.1 2) 2.1 CGI-TD responders, % 1) 18.8 2) 73.7	NR

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Bari, *Neurology*, 2016 [44] **KINECT extension** **POSTER ABSTRACT**	Open label extension (OLE) followed a phase II, double-blind, placebo-controlled period (DBPC) DBPC: 6 weeks OLE: 6 weeks	DBPC 1) Placebo (n=54) 2) VBZ (n=53) OLE 1) VBZ (n=89) Open label extension subjects received once-daily VBZ 50 mg for 6 weeks	NR	Mean age (SD) 55.1 (10.5) Male, n (%) 71 (66.4) Mean age at TD diagnosis (SD) 47.6 (11.8) Mean video AIMS score (SD) 12.3 (5.1)	*Week 12* AIMS change from baseline -5.9 ≥50% AIMS improvement, % 54 CGI-TD mean score (SD) 2.5 (0.9) CGI-TD responders, % 61 *Week 16 (after washout)* CGI-TD responders, % 29	*Week 12* Any treatment emergent AE, n (%) 42 (41.2) Serious AEs, n (%) 0 Suicidal ideation, n 3

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Deutetrabenazine (DTBZ)						
Anderson *Lancet Psychiatry* 2017 [39] **AIM-TD** **Fair**	Phase III, RCT, double-blind, placebo-controlled, parallel-group study 75 sites across USA and Europe 12 Weeks	*Safety population* 1) PBO (n=72) 2) DTBZ, 12mg/d (n=74) 3) DTBZ, 24mg/d (n=73) 4) DTBZ, 36mg/d (n=74) Treatments were administered in 2 divided doses, approximately 10 hrs apart; dose escalation through wk 4 for 24mg and 36 mg, weekly increments of 6 mg/d starting from 12 mg/d; maintenance period for 8 wks; dose could be decreased by 6 mg/d d/t AEs	Inclusion: 18-80 years old with TD diagnosis ≥3 months prior to screening (≥1 month in patients >60); AIMS score ≥6 at screening and baseline; history of DRA use for ≥3 months; stable psychiatric illness on stable medication for ≥30 days; stable antidepressant dose for ≥45 days Exclusion: Neurological condition other than TD; history of suicidal ideation ≤6 months of screening; score ≥11 on the depression subscale of HADS at screening; use of TBZ, reserpine or strong anticholinergic within 30 days	*Safety population* Mean age (SD) 1) 54.2 (12.1) 2) 57 (10) 3) 55.6 (11.3) 4) 58.3 (11.6) Mean duration of TD, yrs (SD) 1) 5.8 (0.7) 2) 5.8 (0.7) 3) 4.4 (0.9) 4) 6.2 (0.7) DRA at baseline, n (%) 1) 45 (77.6) 2) 45 (75.0) 3) 37 (75.6) 4) 35 (63.6) Schizophrenia /mood disorders (%) 1) 58/42 2) 53/47 3) 68/32 4) 59/41	*Week 12: mITT population* LS Mean Change in AIMS score 1) -1.4 2) -2.1 (NS) 3) -3.2 (p=0.003 vs PBO) 4) -3.3 (p=0.001 vs PBO) ≥50% improvement in AIMS score, % 1) 12 2) 13 (NS) 3) 35 (p=0.005 vs. PBO) 4) 33 (p=0.007 vs. PBO) CGIC treatment success, % 1) 26 2) 28 (NS) 3) 49 (p=0.014 vs PBO) 4) 44 (p=0.059 vs PBO) PGIC treatment success, % 1) 31 2) 23 3) 45 4) 40 All NS vs. PBO	*Week 12: safety population* Serious AEs, n (%) 1) 4 (6) 2) 2 (3) 3) 6 (8) 4) 4 (5) Withdrawal due to AEs, n (%) 1) 2 (3) 2) 4 (5) 3) 2 (3) 4) 3 (4) Nervous system disorders, n (%) 1) 10 (14) 2) 6 (8) 3) 6 (8) 4) 16 (22) Psychiatric disorders, n (%) 1) 7 (10) 2) 10 (14) 3) 7 (10) 4) 9 (12)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Jimenez-Shahed *AAN* 2017a [109] **AIM-TD** **POSTER ABSTRACT** ***See Anderson Lancet Psychiatry 2017*** [39]	*See Anderson Lancet Psychiatry 2017* [39]	*mITT population* 1) PBO (n=58) 2) DTBZ, 12mg/d (n=60) 3) DTBZ, 24mg/d (n=49) 4) DTBZ, 36mg/d (n=55) For DTBZ 24 mg/d or 36 mg/d, DTBZ was started at 12 mg/d and titrated up during a 4-week dose escalation phase	*See Anderson Lancet Psychiatry 2017* [39]	*mITT population* Mean age (SE) 1) 54.3 (1.6) 2) 57.9 (1.2) 3) 56.3 (1.5) 4) 59.7 (1.4) Duration of TD, yrs (SE) 1) 5.8 (0.7) 2) 5.8 (0.7) 3) 4.4 (0.9) 4) 6.2 (0.7) Receiving DRA at baseline, n (%) 1) 45 (77.6) 2) 45 (75.0) 3) 37 (75.6) 4) 35 (63.6) Psychotic disorders, n (%) 1) 34 (58.6) 2) 33 (55.) 3) 31 (63.3) 4) 30 (54.5) *See Anderson Lancet Psychiatry 2017* [39]	*Week 12: by baseline DRA use* LS Mean Change in AIMS score *Not receiving DRA* 1) 0 2) -2.4 (p=0.048) 3) -3.1 (p=00.013) 4) -3.1 (p=0.006) *Receiving DRA* 1) -1.7 2) -2.0 3) -3.2 (p=0.036) 4) -3.4 (p=0.017) CGIC treatment success, % *Not receiving DRA* 1) 8 2) 27 3) 58; OR: 16.8 (p=0.011) 4) 60; OR: 18.0 (p=0.004) *Receiving DRA* 1) 31 2) 29 3) 46 4) 34	*See Anderson Lancet Psychiatry 2017* [39]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Jimenez-Shahed AAN 2017b [51] **AIM-TD** **POSTER ABSTRACT** ***See Anderson Lancet Psychiatry 2017*** [39]	*See Anderson Lancet Psychiatry 2017* [39]	1) PBO (n=58) 2) DTBZ, 12mg/d (n=60) 3) DTBZ, 24mg/d (n=49) 4) DTBZ, 36mg/d (n=55) For DTBZ 24 mg/d or 36 mg/d, DTBZ was started at 12 mg/d and titrated up during a 4-week dose escalation phase	*See Anderson Lancet Psychiatry 2017* [39]	*See Anderson Lancet Psychiatry 2017* [39] *& Jimenez-Shahed AAN 2017a* [51]	*Week 12* Odds ratio of AIMS Responder Rates vs. PBO *DTBZ 24mg/d* 50%: 3.96 (p<0.01) 70%: 7.92 (p<0.05) 80%: 1.22 (NS) 90%: Odds was not higher DTBZ 36 mg/d 50%: 3.80 (p<0.01) 70%: 10.76 (p<0.05) 80%: 8.06 (p<0.05) 90%: 3.15 (NS)	*See Anderson Lancet Psychiatry 2017* [39]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Anderson APA 2017a [39] **AIM-TD** ***See Anderson Lancet Psychiatry 2017*** [39]	*See Anderson Lancet Psychiatry 2017* [39]	1) PBO (n=58) 2) DTBZ, 12mg/d (n=60) 3) DTBZ, 24mg/d (n=49) 4) DTBZ, 36mg/d (n=55) Treatments were administered in 2 divided doses, approximately 10 hrs apart; dose escalation through wk 4 for 24mg and 36 mg, weekly increments of 6 mg/d starting from 12 mg/d; maintenance period for 8 wks; dose could be decreased by 6 mg/d d/t AEs	*See Anderson Lancet Psychiatry 2017* [39]	*See Anderson Lancet Psychiatry 2017* [39] *& Jimenez-Shahed AAN 2017a* [51]	*Week 12* Mean mCDQ-24 change (LS) vs placebo 3) -3.5* 4) -4.4* Improvement in QoL by Baseline DRA use *Not receiving DRA* *Vs placebo* 3) -10.1* 4) -6.4* *Receiving DRA* *Vs placebo* *3)* No reduction 4) -2.6 Improvement in QoL by underlying psychiatric conditions *Bipolar:* 3) -5.7* 4) -4.3* *Schizophrenia:* 3) -1 (approx.) 4) -3.2* *clinically meaningful	*See Anderson Lancet Psychiatry 2017* [39]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Fernandez *Neurology* 2017 [40] **ARM-TD** **Fair**	RCT, double-blind, placebo-controlled, multicenter trial 46 sites in the United States and Europe 12 weeks	1) DTBZ (n=58) 2) PBO (n=59) Tablets are taken orally twice daily for 12 weeks; DTBZ intervention includes a dose titration period and maintenance period DTBZ was started at 12mg/d (6 mg twice daily) and titrated weekly by 6 mg/d for up to 6 weeks until adequate dyskinesia control; the dose would increase until a significant AE occurred or the max dose (48 mg/d) was reached; this was followed by a 6 week maintenance and 1 week washout period	Inclusion: DRA use for ≥3 months; clinical diagnosis of TD for ≥3 months prior to screening; patients with psychiatric diagnosis must be stable and have no change in psychoactive medications; have a mental health provider and doesn't anticipate any changes to treatment in next 3 months; AIMS score ≥6 Exclusion: Receiving TD treatment medication; have another neurological condition besides TD; history of substance/alcohol abuse in last 12 months; have a serious untreated/undertreated psychiatric illness; have unstable or serious medical illness; known allergy to any component of SD-809	Mean age (SD) 1) 55.9 (9.8) 2) 53.3 (10.6) White, n (%) 1) 37 (63.8) 2) 44 (74.6) Duration of TD, mo 1) 72.6 (81.7) 2) 76.8 (82.1) DRA use at baseline, n (%) 1) 45 (77.6) 2) 49 (83.1) Baseline AIMS score, items 1—7 (SD) 1) 9.6 (4.1) 2) 9.6 (3.8) Psychiatric disorders *Schizophrenia/ Schizoaffective disorder* 1) 29 (50)/ 11 (19) 2) 29 (49.2)/ 11 (18.6) *Bipolar disorder* 1) 12 (20.7) 2) 15 (25.4)	*Week 12* LS Mean AIMS change (SE) 1) -3.0 (0.45) 2) -1.6 (0.46) P=0.019 95% CI: -2.6 to -0.2 CGIC treatment success, % 1) 48.2 2) 40.4 PGIC treatment success, % 1) 42.9 2) 29.8	*Week 12* Serious AE, n (%) 1) 3 (5.2) 2) 5 (8.5) Treatment-related AEs, n (%) 1) 28 (48.3) 2) 21 (35.6) AEs leading to dose reduction, n (%) 1) 6 (10.3) 2) 3 (5.1) AEs leading to dose suspension, n (%) 1) 3 (5.2) 2) 5 (8.5) AEs leading to discontinuation, n (%) 1) 1 (1.7) 2) 2 (3.4) Somnolence, n (%) 1) 8(13.8) 2) 6 (10.2)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Anderson APA 2017b [49] **AIM-TD and ARM-TD** **POSTER ABSTRACT** ***See Anderson Lancet Psychiatry 2017*** [39] ***See Fernandez Neurology 2017*** [52]	Open-label, single-arm, long-term safety study after washout of treatment in the parent study Data cutoff: June 30, 2016	1) Prior PBO (n=102) 2) Prior DTBZ (n=202) Total N=304 N at week 54=78 Patients were rolled over from parent study after completing 1-week washout and week 13 evaluation Started at 12mg/d and titrated once a week until adequate dyskinesia control was achieved, a clinically significant AE occurred (investigator determined if dose reduction or suspension was necessary), or the max dose of 48 mg/d was reached Study was ongoing through week 158	Inclusion: Successful completion of either ARM-TD or AIM-TD study; adults with clinical diagnosis of TD for ≥3 months prior to screening; history of DRA use for ≥3 months (≥1 months if ≥60 yrs); if psychiatric illness must be stable with no psychoactive medication change for ≥30 days (antidepressants ≥45 days) before screening Exclusion: Neurological condition other than TD; history of suicidal ideation/behavior w/i 6 months of screening; score ≥11 on the depression subscale of the Hospital Anxiety and Depression Scale at screening or baseline	Mean age (SE) 1) 54.3 (1.1) 2) 57.1 (0.7) Female, n (%) 1) 55 (53.9) 2) 114 (5.4) White, n (%) 1) 80 (78.4) 2) 156 (77.2) Mean TD duration, yrs (SE) 1) 5.9 (0.6) 2) 5.6 (0.4) Receiving DRA at baseline, n (%) 1) 80 (78.4) 2) 147 (72.8) Psychotic Disorders, n (%) 1) 62 (60.8) 2) 122 (60.4) Mood disorders, n (%) 1) 39 (38.2) 2) 80 (39.6)	*Week 54* Mean AIMS score improvement (SE) 1) -4.6 (0.77) 2) -5.4 (0.69) Patients who were "Much Improved" or "Very Much Improved" by the CGIC (receiving DTBZ) -At week 6: 58% -AT week 54: 72% By the PGIC scale (receiving DTBZ) -at week 6: 53% -at week 54: 59%	*See* Anderson APA 2017d [110]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Anderson APA 2017c [50] **ARM-TD and AIM-TD** ***See Anderson Lancet Psychiatry 2017*** [39] ***See Fernandez Neurology 2017*** [52]	Open-label, single-arm, long-term safety study after washout of treatment in the parent study Data cutoff: June 30, 2016	1) Prior PBO (n=102) 2) Prior DTBZ (n=202) Total N=304 Patients were rolled over from parent study after completing 1-week washout and week 13 evaluation Started at 12mg/d and titrated once a week until adequate dyskinesia control was achieved, a clinically significant AE occurred (investigator determined if dose reduction or suspension was necessary), or the max dose of 48 mg/d was reached Study was ongoing through week 158	*See* Anderson APA 2017c [49]	*See* Anderson APA 2017c [49]	*See Anderson* Anderson APA 2017c [49]	*Week 54 Exposure-Adjusted Incidence Rate of Patients with AEs (Number of Patients/Patient-year)* Serious AEs: 0.14 (29/202.6) Treatment-related AEs: 0.67 (101/150.1) AEs leading to discontinuation: 0.08 (18/212.4) Nervous system: *Somnolence:* 0.11 (22/201.5) *Headache* 0.10 (21/200.3) Psychiatric disorders *Depression:* 0.11 (22/206.1) *Anxiety:* 0.12 (24/201.5)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Fernandez APA 2017a [52] **AIM-TD and ARM-TD** **POSTER ABSTRACT** ***See Anderson Lancet Psychiatry 2017*** [39] ***See Fernandez Neurology 2017*** [52]	ARM-TD: Phase II/III, RCT, double-blind, placebo-controlled study 12 weeks AIM-TD: Phase III, RCT, double-blind, placebo-controlled, parallel-group study 12 Weeks	*In Pooled Efficacy Population* 1) PBO (n=107) 2) DTBZ (n=152) 152 patients received DTBZ in ARM-TD or 24 or 36mg/d DTBZ in AIM-TD	*See Anderson Lancet Psychiatry 2017* [39] & *Fernandez Neurology 2017* [52]	*In Pooled Efficacy Population* Mean age (SD) 1) 54.5 (10.9) 2) 57.9 (9.9) Female, n (%) 1) 55 (51.4) 2) 84 (55.3) Caucasian, n (%) 1) 84 (78.5) 2) 115 (75.7) Duration of TD, yrs (SD) 1) 6.0 (6.1) 2) 5.6 (6.1) Psychotic disorders, n (%) 1) 68 (63.6) 2) 93 (61.2) Mood disorders, n (%) 1) 38 (35.5) 2) 59 (38.8)	*In Pooled Efficacy Population* *Week 12* Patients "Much Improved" or "Very Much Improved" on the CGIC 1) 30 2) 48 OR: 2.12, p=0.005	*Pooled Safety Population* *Week 12* Serious AE, n (%) 1) 9 (6.9) 2) 13 (6.3) Treatment-related AEs, n (%) 1) 40 (30.5) 2) 57 (27.8) AE leading to discontinuation, n (%) 1) 4 (3.1) 2) 6 (2.9) AE leading to dose reduction, n (%) 1) 3 (2.3) 2) 10 (4.9) Deaths, n (%) 1) 0 2) 2 (1.0) Common AE>3% Somnolence, n (%) 1) 9 (6.9) 2) 12 (5.9) Headache, n (%) 1) 10 (7.6) 2) 10 (4.9)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Fernandez APA 2017b [53] **AIM-TD and ARM-TD** ***See Anderson Lancet Psychiatry 2017*** [39] ***See Fernandez Neurology 2017*** [52]	ARM-TD: Phase II/III, RCT, double-blind, placebo-controlled study 12 weeks AIM-TD: Phase III, RCT, double-blind, placebo-controlled, parallel-group study 12 Weeks	*In Pooled Efficacy Population* 1) PBO (n=107) 2) DTBZ (n=152) 152 patients received DTBZ in ARM-TD or 24 or 36mg/d DTBZ in AIM-TD	*See Anderson Lancet Psychiatry 2017* [39] & *Fernandez Neurology 2017* [52]	Fernandez APA 2017a [52]	*Week 12* Patients "Much Improved" or "Very Much Improved" on PGIC *Overall* 1) 30 2) 43 OR: 1.81; p=0.026 *With Underlying Mood Disorder* 1) 26/38 2) 47/59 OR: 2.76; p=0.028 *With Underlying Psychotic Disorder* 1) 32/68 2) 41/93 OR: 1.38; p=0.342	Fernandez APA 2017a [52]

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Tetrabenazine (TBZ)						
Miguel R, Ther Adv Neurol Disord. 2017 [62] **Poor**	Retrospective observational study Duration of TBZ treatment, months Median: 40 Range: 1 – 239 Study population: Patients with hyperkinetic movement (e.g. Tardive OM dyskinesia, TS, HD, chorea)	TBZ Initial TBZ dose was 12.5mg and maximum daily dose was 37.5mg. *Tardive Oromandibular Dyskinesia* (n=35) *Other Tardive Syndrome** (n=22) Other TS population is defined as patients with dystonia, chorea, akathisia and tic manifestations. *Total population on TBZ (n=111), subpopulations of focus mentioned above	Inclusions: All consecutive patients with hyperkinetic movement disorders who were under TBZ treatment, observed between 1 January 2006 and 31 December 2015, in Egas Moniz Hospital Outpatient Clinic were included. Exclusion: Insufficient follow up or insufficient compliance information.	Tardive OM dysk patients Males, n (%) 35 (45.7) Age at TBZ onset, yrs 75 Concomitant medication, n (%) 4 (11.4) Duration of motor symptoms before TBZ onset, m 12 Other TS patients Males, n (%) 22 (13.4) Age at TBZ onset, yrs 64.5 Concomitant medication, n (%) 10 (47.6) Duration of motor symptoms before TBZ onset, m 12	*Tardive OM dysk patients* Total responder, (n) 27 Asymptomatic responders, n (%) 3 (11.1) Improvement but symptomatic, n (%) 24 (88.9) Nonresponder, (n) 8 Responder rate, (%) 77.1 *Other TS patients* Total responder, (n) 21 Asymptomatic, n (%) 3 (14.3) Improvement but symptomatic, n (%) 18 (85.7) Nonresponder, (n) 1 Responder rate, (%) 95.5	Tardive OM dys / TS patients AE, n (%) 21 (60)/ 9 (40.9) Parkinsonism, n (%) 14 (40)/ 7 (31.8) Psychiatric disorders, n (%) 3 (8.6)/ 1 (4.5) *Other side effects, n (%) 3 (8.6)/ 2 (9.1) All patients AE, n (%) 52 (48.1) Parkinsonism, n (%) 29 (26.9) Psychiatric disorders, n (%) 14 (13) *Other side effects, n (%) 11 (10.2) *Somnolence, rash, mental confusion

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Kazamatsuri H Arch Gen Psychiat 1972 [41] **Poor**	Non-randomized cross over study Duration of follow-up: 6 weeks At four weeks, all patients were placed on placebo, and all the neuroleptic drugs were completely withdrawn for four weeks.	PBO (4 weeks) TBZ (6 weeks) 50mg daily for first 2 weeks, 100mg daily for next 2 weeks, 100 or 150mg daily for last 2 weeks. PBO (2 weeks) Total N (n = 24)	Inclusions: All patients on the chronic wards of Boston State Hospital with oral dyskinesia. Exclusion: Patients with borderline cases not showing clear oral dyskinesia were eliminated.	Males, (n) 13 Mean age, yrs (range) 55 (30-81) Mean length of psychiatric hospitalization, yrs (range) 28.8 (10-52) Psychiatric diagnoses, n Chronic schizophrenia: 17 Chronic brain syndrome: 4 Mental deficiency: 3 Baseline oral dyskinesia mean frequency/minute: 29.6	Oral dyskinesia Mean frequency/ minute: 1) 30 (after 4 weeks of PBO) 2) 10.8 (after 6 weeks of PBO) p value 2 vs.1 <0.0005 Efficacy of TBZ compared with PBO, n (%) At 2 weeks: 100% symptoms disappearance: 3 (12.5) 50% or more reduction: 9 (37.5) At 4 weeks 100% symptoms disappearance: 4 (16.7) 50% or more reduction: 13 (54) At 6 weeks 100% symptoms disappearance: 8 (33.3) 50% or more reduction: 9 (58.3)	TBZ related AE, n (%) Severe malaise, 2 (8.3) Psychotic exacerbation 1 (4.2)

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Jankovic J Neurology 1988 [57] **Fair**	Single-Arm, Open Label, Long term Study Mean follow-up, m (range): 18 (1 to 80) Study population: Patients with hyperkinetic movement (e.g. TD, Huntington's disease, Tourette's syndrome)	*1) TBZ* Tardive dyskinesia (n=44) patients started on 25mg/d, dose increase of 25mg/d until 100 mg was reached/patient noted adverse side effects. Tardive dystonia (n=15) patients started on 25mg/d and dose increase of 25mg/d until 100 mg was reached/pt noted adverse side effects *Total population on TBZ (n=217), subpopulations of focus mentioned above	Inclusions: Patients with movement disorder interfering with activities of daily living or with their job; and must have failed all other conventionally used medication (e.g. patients with TD must have been first treated with reserpine).	Tardive dyskinesia Mean age at onset of movement disorder, yrs (range) 60.4 (34-83) Mean duration of symptoms before TBZ, yrs (range) 2.5 (0.5-11) Mean duration of TBZ therapy, m 21.1 (0.12-79) Tardive dystonia Age at onset of movement disorder, yrs 36.1 (2-67) Mean duration of symptoms before TBZ, yrs (range) 3.9 (0.2-21) Mean duration of TBZ therapy, m 15.7 (0.75-57)	*Global response scale (1 = marked improvement, 2 = moderate improvement, 3 = fair improvement, 4 = no response, 5 = worsening) <u>*Tardive dyskinesia*</u> Average response to TBZ Scale* of 1-5: 2.3 Response to TBZ on a scale of 1-5, n (%) Marked improvement, 6 (14) Moderate improvement, 25 (57) Fair improvement, 11 (25) No response, 1 (2) Worsening, 1 (2) <u>*Tardive dystonia*</u> Average response to TBZ Scale* of 1-5: 2.6 Response to TBZ Scale of 1-5, n Marked improvement, 0 Moderate improvement, 8 Fair improvement, 6 No response, 1 Worsening, 0	*TBZ related AE, n Parkinsonism, 53 Drowsiness/ fatigue, 28 Depression, 23 Anxiety, 16 Insomnia, 11 Akathisia, 10 *(n=217) – Total population on TBZ

<table>
<tr><th>Author & Year of Publication (Trial Name) Quality rating</th><th>Study Design and Duration of Follow-up</th><th>Interventions (n) & Dosing Schedule</th><th>Inclusion and Exclusion Criteria</th><th>Patient Characteristics</th><th>Outcomes</th><th>Harms</th></tr>
<tr>
<td>Jankovic J Neurology 1997 [58]

Fair</td>
<td>Single-Arm, Open Label, Long term Study

Mean follow-up, m (range): 28.9 (0.25 to 180)</td>
<td>1) TBZ
Treatment regimen was modified during the study: Patients seen before 1991 were started at 25 mg/d, with a dose increase of 25 mg/d until a total dose of 150-200 mg/d was reached or observed side effect. Patients seen since 1991 were maintained at a dose of 25 to 75 mg/d.

Tardive dyskinesia (n=93)
Tardive dystonia (n=82)

*Total population on TBZ (n=400), subpopulations of focus mentioned above</td>
<td>Inclusions: Patients with various hyperkinetic movement disorders

Exclusions: A single evaluation without a follow-up visit, insufficient information about response, discontinued TBZ within the first 2 weeks, already on TBZ when first seen in the clinic, or noncompliant (n=126/526)</td>
<td>Tardive dyskinesia (n=93)
Males; Females, n
22; 71
Mean age at onset of movement disorder, yrs
63 (7-86)
Maximum daily dose, mg
96 (25-400)
Duration of TBZ therapy, m
35 (0.3-171)

Tardive dystonia (n=82)
Males; Females, n
22; 60
Mean age at onset of movement disorder, yrs
45 (2-72)
Maximum daily dose, mg
125 (38-400)
Duration of TBZ therapy, m
32 (0.75-180)

All patients
Mean duration of symptoms before TBZ, yrs
6.6 (3 weeks to 48 years)</td>
<td>*Global response scale (1 = marked improvement, 2 = moderate improvement, 3 = fair improvement, 4 = no response, 5 = worsening)
Tardive dyskinesia
Response to TBZ Scale* of 1-5, n (%)
<table>
<tr><th></th><th>1) IR†</th><th>2) LR†</th></tr>
<tr><td>1</td><td>83 (89.24)</td><td>79 (84.9)</td></tr>
<tr><td>2</td><td>4 (4.3)</td><td>7 (7.4)</td></tr>
<tr><td>3</td><td>2 (2.1)</td><td>1 (1.1)</td></tr>
<tr><td>4</td><td>4 (4.3)</td><td>5 (5.3)</td></tr>
<tr><td>5</td><td>0</td><td>1 (1.18)</td></tr>
</table>
Tardive dystonia
Response to TBZ Scale* of 1-5, n (%)
<table>
<tr><th></th><th>1) IR†</th><th>2) LR†</th></tr>
<tr><td>1</td><td>66 (80.5)</td><td>60 (73.2)</td></tr>
<tr><td>2</td><td>6 (7.3)</td><td>8 (9.8)</td></tr>
<tr><td>3</td><td>7 (8.5)</td><td>7 (8.5)</td></tr>
<tr><td>4</td><td>3 (3.7)</td><td>3 (3.7)</td></tr>
<tr><td>5</td><td>0</td><td>0</td></tr>
</table>
†IR – Initial Response (3 to 6 months after starting TBZ)
†LR- Response at last visit</td>
<td>TBZ related AE, n (%)

Drowsiness/ fatigue 146 (36.5)

Parkinsonism 114 (28.5)

Depression, 60 (15.0)

Insomnia, 44 (11.0)

Anxiety, 41 (10.3)

Akathisia, 38 (9.5)

*Total population on TBZ (n=400)</td>
</tr>
</table>

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Paleacu D Clin Neuropharmacol 2004 [61] **Poor**	Retrospective observational study Mean follow-up time, m 22	*1) TBZ* *Tardive dyskinesia* (n=17) starting dose of 12.5 mg BID, dose increase of 25 mg once a week to a target dose of 150 mg/d in 2 or 3 divided doses daily. *Total population on TBZ (n=118), subpopulation of focus mentioned above	Inclusions: Patients with a hyperkinetic movement disorder including dystonia, Huntington disease (HD) or other choreas, tardive dyskinesia (TD) or akathisia, and Tourette syndrome.	*Total Population on TBZ* Mean Age, yrs 48.8 Males, n 48 Mean disease duration, m 93 Mean TBZ dose, mg/d (range) 76.2 (25-175) Mean CGIC score +1 (mild improvement)	*Tardive dyskinesia* CGIC Scale* of 1-5, n **No Response or Worsening** **0**: 4 **-1**: 1 **-2**: None **-3**: None **Improved** **1**: 3 **2**: 6 **3**: 1 No data: 2 *Clinical Global Impression of Change (CGIC) −3 was marked worsening, −2 moderate worsening, −1 mild worsening, 0 no change, +1 was mild improvement, +2 moderate improvement, and +3 marked improvement	Total Population on TBZ TBZ related AE, n *Somnolence/ Weakness/Apathy, 5 Depression, 2 Parkinsonism, 6 Acute akathisia, 1 Tardive dyskinesia TBZ related AE, n *Somnolence/ Weakness/Apathy, 2 Worsening of underlying condition, 2 *4/7 patients among this group discontinued treatment.

<table>
<tr><th>Author & Year of Publication (Trial Name) Quality rating</th><th>Study Design and Duration of Follow-up</th><th>Interventions (n) & Dosing Schedule</th><th>Inclusion and Exclusion Criteria</th><th>Patient Characteristics</th><th>Outcomes</th><th>Harms</th></tr>
<tr>
<td>Kenney C Mov Disoders 2007 [60]

Poor</td>
<td>Retrospective observational study

Time frame – 1997-2004

Mean follow-up time: 2.3 -3.4 years</td>
<td>1) TBZ

January 1997- 2004 (n=448)
Patients treated with TBZ 50-75 mg/d; 18.2% of patients required doses greater than 75 mg/d (range: 12.5–300 mg/d).</td>
<td>Inclusions: Patients with involuntary movements that were troublesome or disabling despite optimal conventional therapy.</td>
<td>Tardive dyskinesia (n=149)
Mean age at onset of movement disorder, yrs
59.8
TBZ treatment duration, yrs
2.5
Baseline severity (prior to TBZ), %
Mild, 1.3; Moderate, 40.3; Severe, 50.3; Disabling, 8.1

Tardive dystonia (n=132)
Mean age at onset of movement disorder, yrs
44.6
TBZ treatment duration, yrs
3.0
Baseline severity (prior to TBZ), %
Mild, 0; Moderate, 37.4; Severe, 45.8; Disabling, 16.8

All patients (n=448)
Mean daily dose of TBZ: 60.4 mg/d</td>
<td>Tardive Dyskinesia
Efficacy response to TBZ Scale* of 1-5, %
<table>
<tr><th></th><th>Initial Response</th><th>Last response</th></tr>
<tr><td>1</td><td>65%</td><td>71%</td></tr>
<tr><td>2</td><td>18%</td><td>13%</td></tr>
<tr><td>3</td><td>6%</td><td>4%</td></tr>
<tr><td>4</td><td>5%</td><td>5%</td></tr>
<tr><td>5</td><td>3%</td><td>3%</td></tr>
</table>
Tardive Dystonia
Efficacy response to TBZ Scale* of 1-5, %
<table>
<tr><th></th><th>Initial Response</th><th>Last response</th></tr>
<tr><td>1</td><td>45%</td><td>45%</td></tr>
<tr><td>2</td><td>21%</td><td>23%</td></tr>
<tr><td>3</td><td>12%</td><td>10%</td></tr>
<tr><td>4</td><td>18%</td><td>16%</td></tr>
<tr><td>5</td><td>2%</td><td>3%</td></tr>
</table>
*Modified AIMS score (1 = marked improvement, 2 = moderate improvement, 3 = fair improvement, 4 = no response, 5 = worsening)
(% values in both tables are represented as approximate values from the text).</td>
<td>*TBZ related AE, n (%)

Drowsiness, 112 (25)

Parkinsonism, 69 (15.4)

Depression, 34 (7.6)

Akathisia, 34 (7.6)

Nausea/Vomiting 25 (5.6)

*Total population on TBZ (n=448)</td>
</tr>
</table>

Author & Year of Publication (Trial Name) Quality rating	Study Design and Duration of Follow-up	Interventions (n) & Dosing Schedule	Inclusion and Exclusion Criteria	Patient Characteristics	Outcomes	Harms
Ondo W AM J psych 1999 [59] **Poor**	Randomized single-blind Mean duration of follow-up – 20.3 weeks Patients stopped taking offending medications and any other treatments for tardive dyskinesia at least 30 days before they entered the study.	1) TBZ (n=20) starting dose of 12.5/mg BID and then increased to a maximum of 50/mg TID in weekly increments until satisfactory benefit was experienced or if adverse events became troublesome.	Inclusions: Patients diagnosed with TD who didn't respond to conventional anti-tardive-dyskinesia treatment Patients were required to stop taking offending medications at least 30days before start of study	Females:75% Mean age, yrs: 65.2 Drugs causing TD, n Metoclopramide:7 Haloperidol: 7 Chlorpromazine:2 Perphenazine: 1 Thiethylperazine: 1 Amoxapine: 1 Fluphenazine:1 Mean duration of TD, m 43.7 (2-420) Mean TBZ dose, mg/d (SD) 57.9 (22.8) Mean baseline AIMS score (SD) By blinded raters 17.9 (4.4) Subjective score 9.1 (1.5)	Improvement in AIMS score, Mean change from baseline By blinded raters 9.7 (p<0.001) Subjective 5.5 (p<0.001)	TBZ related AE, n Sedation, 5 Mild Parkinsonism, 5

Appendix H. Comparative Value Supplemental Information

Table H1: Impact Inventory [111]

Sector	Type of Impact	Included in This Analysis from… Perspective?		Notes on Sources
		Health Care Sector	Societal	
Formal Health Care Sector				
Health outcomes	Longevity effects	☒	☒	
	Health-related quality of life effects	☒	☒	
	Adverse events	☒	☒	
Medical costs	Paid by third-party payers	☒	☒	
	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	☒	
	Cost of unpaid lost productivity due to illness	NA	☐	
	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	☐	

Table H2. Drug Cost Inputs

Drug Name, Labeled Dose, Administration Route	Package Size	WAC	Annual WAC	Annual Net Price*	Source
Valbenazine 80mg Oral	-	-	$75,960	$55,451	Manufacturer input[112]
Deutetrabenazine 36mg Oral	12mg 60s ea	$4,932	$90,009	$65,707	Redbook, 2017[113]

*Assumed discount of 27%

Table H3. Non- Drug Cost Inputs

	Base Case Estimate	Source
General Practitioner Visit (CPT Code 99213, Non-Institutional Charge)	$73.93	2017 Physician Fee Schedule[82]
Specialist (Neurologist) Visit (CPT Code 99213, Non-Institutional Charge)	$108.74	2017 Phsycian Fee Schedule[82]

Table H4. Utility Inputs

Parameters	Base Case Value	Reference	Notes
Baseline Utility for Modeled Population	0.82	Wang 2004[84] Calvert 2006[83]	Weighted average of patients with schizophrenia disorders (0.83) and bipolar disorders (0.80).
Utility Decrement from TD	0.095	Lenert 2004[114]	Assumed constant across age and underlying condition.

Table H5. Probabilistic Sensitivity Analysis Inputs

Input Varied	Distribution
Proportion of Placebo Responders (Valbenazine Comparison)[36]	Beta (α = 6.61, β = 69.39, μ = 8.7)
Proportion of Valbenazine Responders[36]	Beta (α = 31.60, β = 47.40 μ = 40.0)
Proportion of Placebo Responders (Deutetrabenazine Comparison)[75]	Beta (α = 7.00, β = 51.00 μ = 12.0)
Proportion of Deuetetrabenazine Responders[75]	Beta (α = 24.50, β = 49.50 μ = 33.1)
Valbenazine Discontinuation Rate[36,77]	Beta (α = 43.70, β = 186.30, μ = 19.0)
Deutetrabenazine Discontinuation Rate[76]	Beta (α = 2, β = 13.4, μ = 13.0)
Disutility with TD	Beta (α = 95, β = 905, μ = 9.5)
Annual Cost of TD	Uniform (min = 183, max = 548, μ = 365.34)
Annual Cost of Valbenazine	Uniform (min = 27,726, max = 83,177, μ = 55,451)
Annual Cost of Deutetrabenazine	Uniform (min = 32,854, max = 98,561, μ = 65,707)

Figure H1. Incremental Cost-Effectiveness Scatterplot of Valbenazine – Placebo

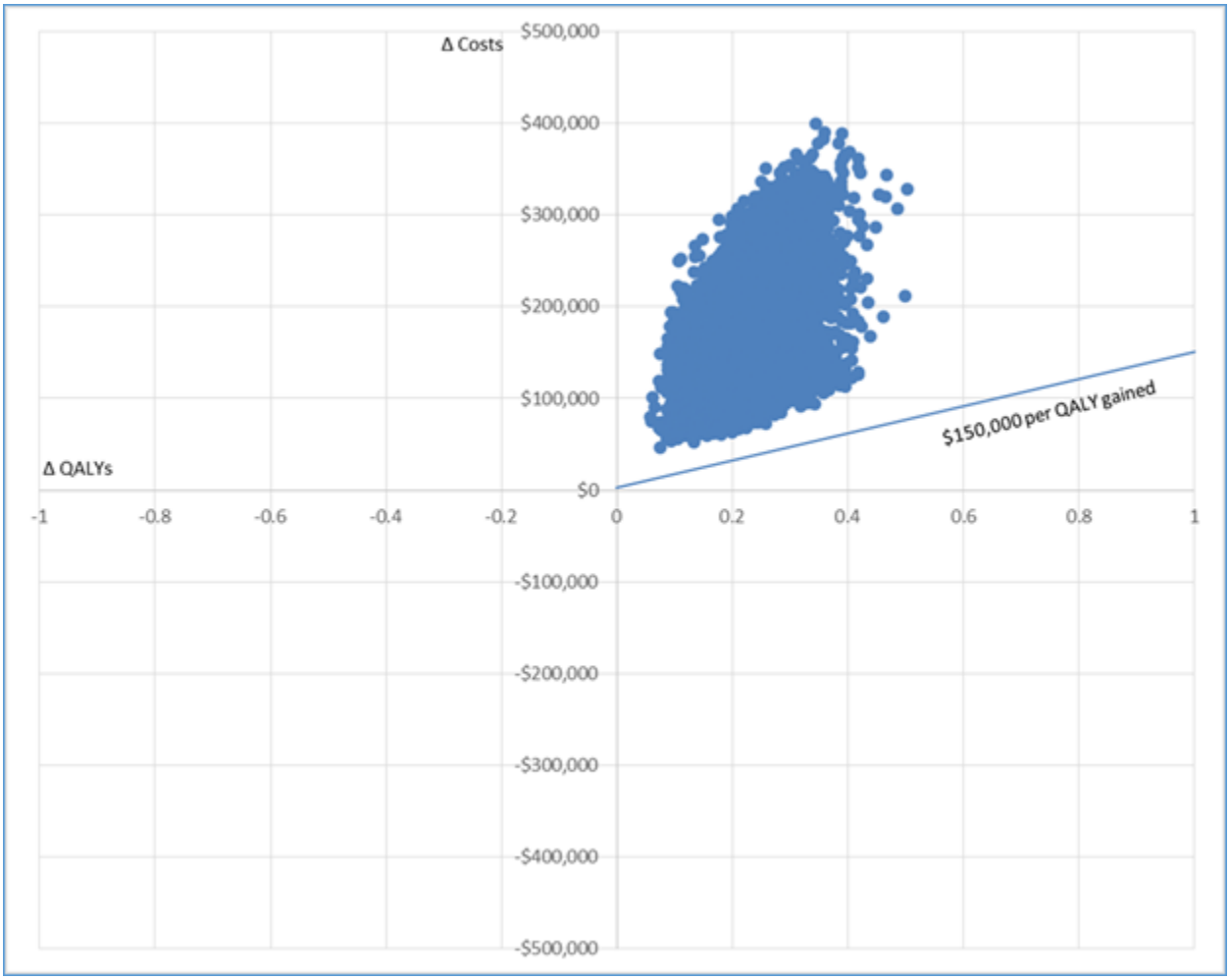
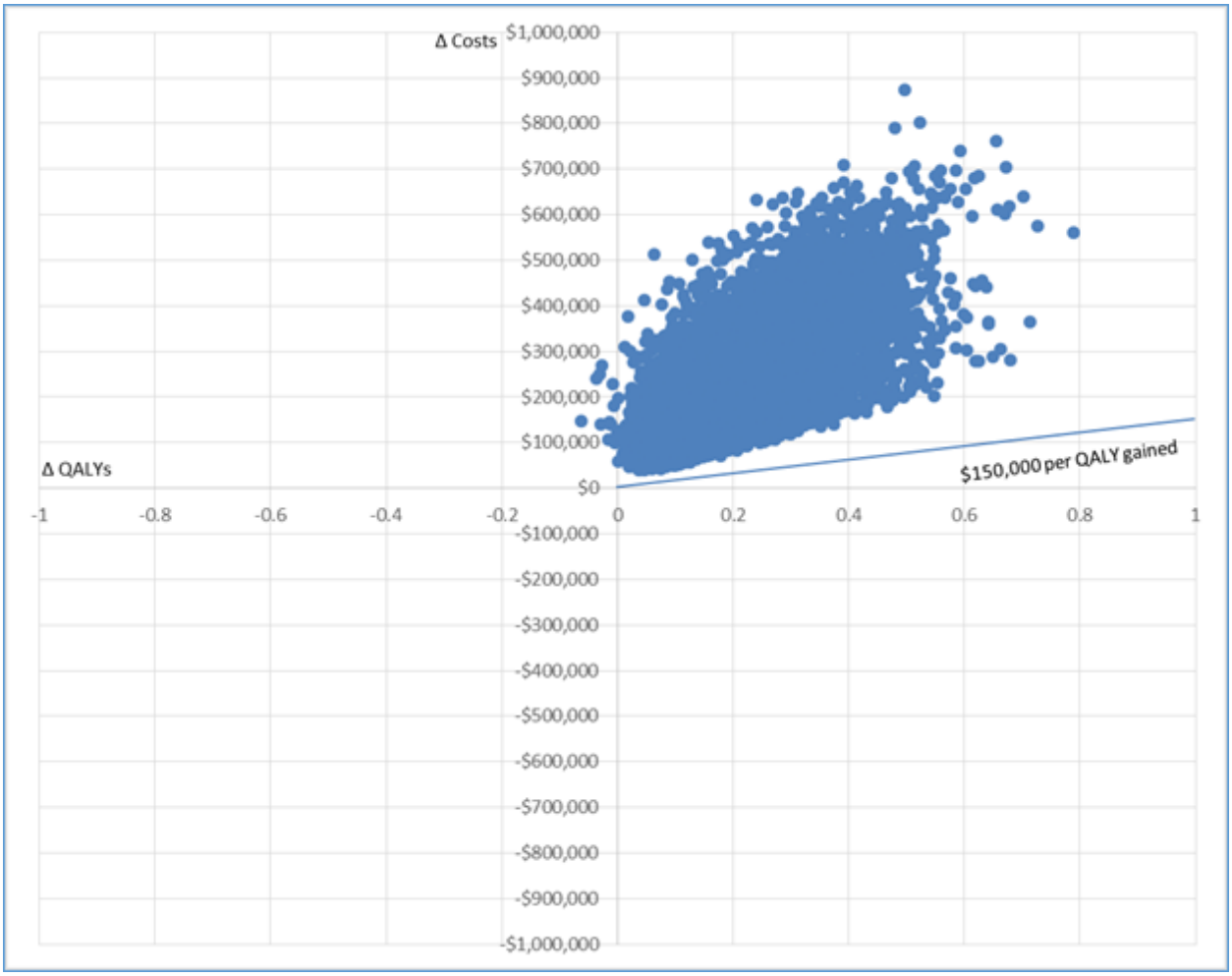


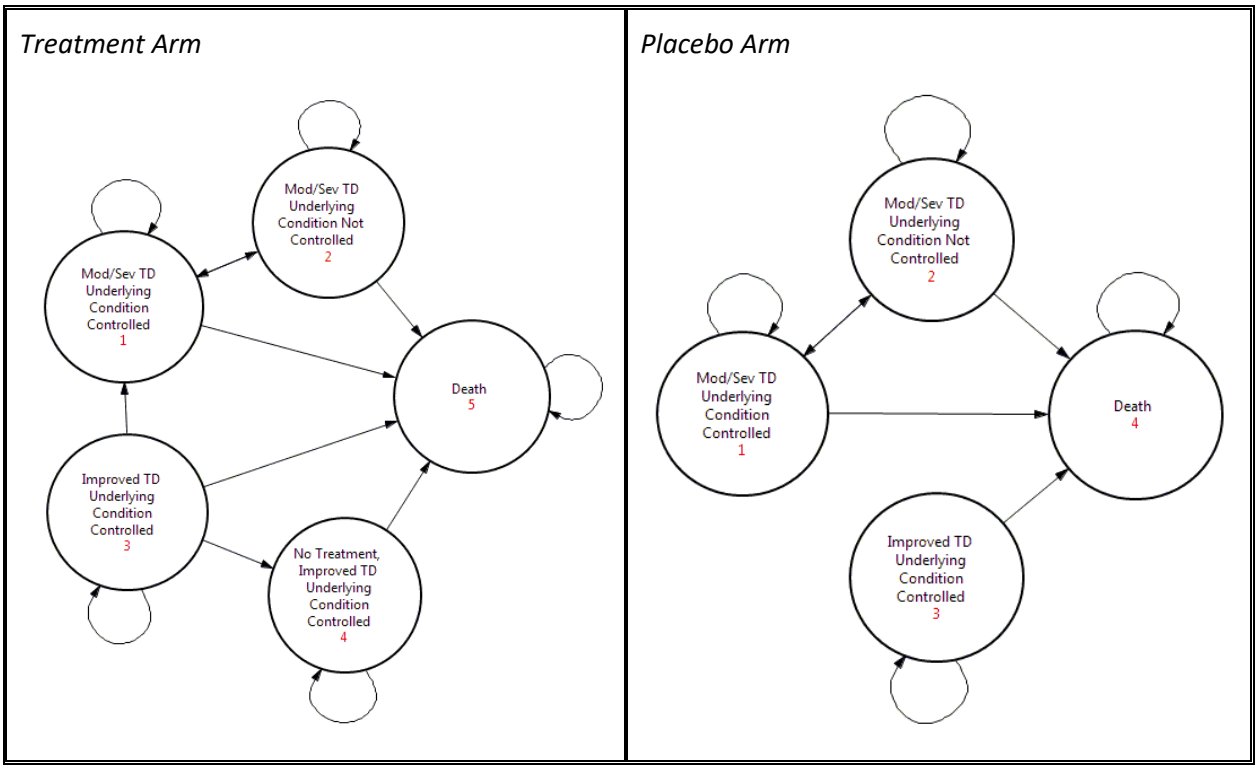
Figure H2. Incremental Cost-Effectiveness Scatterplot of Deutetrabenazine – Placebo



Scenario Analyses

The five main health states for the scenario analyses are as follows: 1) moderate to severe TD with the underlying condition well controlled, 2) moderate to severe TD with the underlying condition poorly controlled, 3) improved TD (on TD treatment) with the underlying condition well controlled, 4) improved TD (without TD treatment) with the underlying condition controlled, and 5) death (Figure H3). It has been hypothesized that improved tardive dyskinesia control results in better management of the underlying conditions. In these scenario analyses, treatment with valbenazine and deutetrabenazine were modeled to improve management of schizophrenia/schizoaffective disorders and bipolar disorders, resulting in additional utility gains and reduced total costs. To model this possibility, patients with moderate to severe TD were given a 10% probability each year of having their underlying condition become poorly controlled. Control of the underlying condition was assumed to be re-established each subsequent cycle. The probability of developing poorly controlled schizophrenia/schizoaffective disorder or bipolar disorder was varied in a threshold analysis from 0-100%.

Figure H3. Scenario Analysis Model Structure



Additional model inputs were utilized in the scenario analyses (Table H5). These additional inputs were obtained from published literature where possible. In particular, the proportion of patients with TD becoming non-adherent to their antipsychotic medication had no associated data in the literature so a potential effect was used. The annual costs of well-controlled and poorly-controlled schizophrenia/schizoaffective and bipolar disorders were taken from the literature and then inflation adjusted to 2016.

Table H6. Additional Model Inputs for the Scenario Analysis Model

Parameters	Value	Reference	Notes
Proportion of Patients with TD Becoming Non-adherent to Antipsychotics	10.0%	Unknown; assumed	Applied each cycle
Annual Cost of Well Controlled Schizophrenia/Schizoaffective Disorders	$22,175	Peng 2011[115]	Inflation-adjusted to Dec. 2016
Annual Cost of Poorly Controlled Schizophrenia/Schizoaffective Disorders	$31,253	Peng 2011[115]	Inflation-adjusted to Dec. 2016
Annual Cost of Well Controlled Bipolar Disorder	$6,568	Calvert 2006[83]	Inflation-adjusted to Dec. 2016
Annual Cost of Poorly Controlled Bipolar Disorder	$10,722	Calvert 2006[83]	Inflation-adjusted to Dec. 2016
Productivity Losses Due to TD	$2,252	Ascher-Svarnum 2008[87]	Based on study of employment differences with TD and median salary in the US in 2016
Disutility Resulting from Poorly Controlled Schizophrenia/Schizoaffective Disorders	-0.27	Wang 2004[84]	
Disutility Resulting from Poorly Controlled Bipolar Disorder	-0.24	Calvert 2006[83]	

Table H7. Incremental Cost-Utility, Alternative Time Horizons

Base Case Versus Placebo	Valbenazine	Deutetrabenazine
	Incremental Cost per QALY	Incremental Cost per QALY
1-year	$828,160	$1,249,854
2-year	$789,651	$1,173,020
5-year	$766,187	$1,125,353