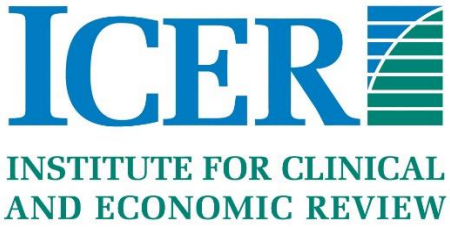
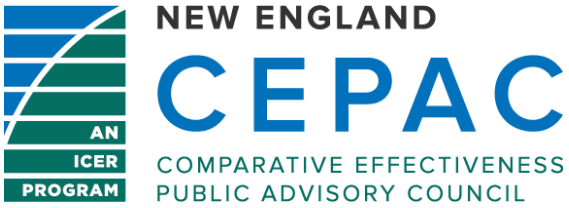


Vesicular Monoamine Transporter 2 Inhibitors for Tardive Dyskinesia: Effectiveness and Value

Response to Public Comment

November 21, 2017

Prepared for:



	Who?	Comments on TD Draft	ICER Response
2	Mental Health America	MHA appreciates that ICER has limited published research on the different aspects of TD to draw from, including limited data on prevalence, disutility, impact on adherence, and impact on employment and related role fulfillment. ICER noted the gaps in evidence well, and included meaningful contextual considerations from a number of sources. ICER also took into account some of this uncertainty in its sensitivity and scenario analyses.	Thank you. We do our best to capture all of the evidence from published and unpublished sources; collect perspectives from expert stakeholders; test key data points in sensitivity analyses; and better understand the many layers of a disease and its impact on patients through scenario analyses.
3	Mental Health America	MHA urges ICER to integrate more of the contextual considerations, as appropriate, into the quantitative sections that are likely most helpful to decision-makers—rather than noting them separately in qualitative sections	The quantitative evaluation is only one aspect of our report. At our meeting, we will discuss in depth the comparative clinical effectiveness, other benefits, and contextual considerations which feature prominently in our evaluation and in our value framework.
4	Mental Health America	MHA recommends to ICER, in particular, that it helps readers understand the utility decrement sensitivity analyses in the context of the study it came from, and the stakeholder input offered. Much of the feedback from stakeholders revolved around the impacts of TD on the lives of individuals, especially those aspects that may be difficult to capture in clinical research. The utility decrement (UD) estimate comes from a well-conducted study, but is derived from this single study that was determining the UD, to answer a specific question about the different perspectives of	We have clarified the description of how utilities were derived in this study and acknowledge the limitations. Given these limitations, we award a full benefit (i.e., as if TD symptoms were completely eliminated) for patients who achieved only a 50% or greater improvement on the AIMS score, which we believed to be a generous benefit. This aspect of the methods has also been expanded in the methods description.

	Who?	Comments on TD Draft	ICER Response
		patients, family members, and providers on the treatment options that existed at the time.	
5	Mental Health America	It is possible that the UD questions may have been asked differently in the context of different study aims, or if conducted today. The sensitivity analyses recognize this possible limitation, as do the sections on stakeholder input. It would be helpful for ICER to note more directly how to understand the sensitivity analyses in the context of the stakeholder input, and the Depression Bipolar Support Alliance (DBSA) survey in particular, even if that is to say that there is no relationship between the two. This could help more meaningfully integrate the quantitative and qualitative sections of the report.	Utilities are typically solicited using very specific questions and methods. Even with these rigorous methods, it is possible that using different solicitation techniques and drawing from different populations can result in different utilities. Often, but not always, using scenarios with healthy volunteers (instead of those with the condition) result in greater utility decrements than those obtained from patients with the condition, because patients adapt to living with their illnesses. We do not know if this is the case for TD. In the absence of better estimates, we believe that the estimates used in our analysis are the best available. At the same time, we acknowledged that these estimates may be biased and therefore have conducted extensive sensitivity analyses. The report has been updated to be more transparent regarding how the utility values used in the model were solicited and the limitations of this method. We have also included a comparison of mental health-related and other medical conditions that produce similar disutilities to place these estimates in context. The DBSA survey was conducted to get a better understanding of the impact of TD and potential treatments on patients' lives. However, this survey was not conducted using standard survey methodology. Nor was it designed to quantitatively assess utility estimates. As such these results, while informative, cannot and should not be included in a quantitative assessment of patient quality of life.
6	Mental Health America	MHA also reiterates its previous request for ICER to conduct three additional sensitivity/scenario analyses. Even though some of these points are acknowledged in the qualitative section, they are not integrated into the quantitative sections –which are likely of the greatest use to decision-makers: 1. Benefits of disenrollment from public payers. Medicaid and disability Medicare are the largest payers of behavioral health services in the United States. Social determinants from poverty and disability can	We agree that these analyses would be informative and considered ways to capture the effects of treating TD on public programs. We found, however, there is currently insufficient information available to be able to develop these scenarios as a quantitative analysis. Without being able to quantify these effects, our analyses would need to include many assumptions and the results would be unreliable.

	Who?	Comments on TD Draft	ICER Response
		lead to behavioral health conditions, and behavioral health conditions create burdens that can lead to poverty and disability. Effective treatment and management of behavioral health conditions, on the other hand, can break this reinforcing cycle and allow individuals to reach a level of participation in community life that allows them to purchase commercial insurance and no longer require public benefits. From a health care payer perspective, this is different than the increases in productivity that ICER currently evaluates. With Medicaid and disability Medicare, increases in productivity beyond a threshold uniquely reduce health care costs for the public payer as the individual disenrolls entirely. Such a scenario analysis would benefit the field. By making such analyses common practice, it can shift the paradigm for how CMS and state Medicaid agencies view costs and benefits – away from trimming health care costs and toward making critical investments that alleviate poverty and disability.	

	Who?	Comments on TD Draft	ICER Response
7	Mental Health America	conduct three additional sensitivity/scenario analyses... 2. Benefits for individuals for family being treated. ICER currently models productivity benefits when feasible, which matter to the employers that contract with the health care payer. In many cases however the treatment is rendered to a family member covered by the employer-sponsored plan, not to the employee themselves. For example, in TD, the spouse or child of a covered individual may receive treatment related to bipolar disorder or schizophrenia. In this case, the productivity benefits accrue indirectly – effectively managed bipolar disorder or schizophrenia in the family member can allow for a more productive individual. Currently, employer benefits managers may believe that such considerations are important in selecting a health plan, but often lack rigorous quantitative methods to incorporate these considerations in cost-benefit determinations. A scenario analysis of how TD treatment for a family member of a working individual would affect that individual's productivity could help employers begin to better integrate the indirect effects of health care on worker productivity in health care purchasing decisions.	We agree that these analyses would be informative. However, there is currently insufficient information available to be able to develop these scenarios as a quantitative analysis. Critical missing information that is needed to conduct this analysis include the proportion of patients needing caregivers as a result of TD (and not the underlying condition) and lost caregiver wages (including time off work and salary). Information would also be required regarding the impact of the TD medications on the ability of caregivers to spend more time working or doing other activities, as it cannot be assumed that a medication that reduces, but does not eliminate TD, would have an impact on the need for a caregiver. Without being able to quantify these effects, our analyses would need to include many assumptions and the results would be unreliable.

	Who?	Comments on TD Draft	ICER Response
8	Mental Health America	Conduct three additional sensitivity/scenario analyses...3. Benefits related to changing social norms. Many individuals experiencing early symptoms of schizophrenia, and even full psychosis, often do not receive effective treatments for months or years. In part, this is due to lack of awareness but it is also sometimes due to the implications of giving a diagnosis of schizophrenia and initiating treatment. The availability of effective treatments for TD may decrease reticence to identify schizophrenia in adolescence and therefore could mean intervening early, because the perceived negative side-effects associated with treatment initiation will be lessened. If people do identify and intervene more readily, this would make the underlying treatments dramatically more effective, as NIMH's Recovery After an Initial Schizophrenia Episode (RAISE) repeatedly finds – the earlier treatment begins, the better the prognosis. Scenario analyses of effects on earlier intervention for decreases in stigma associated with schizophrenia treatment may help understand how the availability of new treatments alters social norms over the long-term, a concept that has broad applicability outside TD treatment as well.	We agree that these analyses would be informative. However, there is not enough information to be able to develop these scenarios as a quantitative analysis. Without being able to quantify these effects, our analyses would be unreliable. Relative to all 3 points above, we note that we did conduct threshold analyses indicating how large improvements in the management of underlying conditions would need to be for the new agents to achieve commonly-accepted cost-effectiveness thresholds. The findings from this analysis are discussed extensively in our report.
9	Teva	Black box warning: Deutetrabenazine (Austedo®, Teva) labeling carries a "black box" warning for depression and suicidality in patients with Huntington's disease. This	We have revised our statements on the "black box" warning throughout the report.

	Who?	Comments on TD Draft	ICER Response
		specific context of “in patients with Huntington’s disease” is missing in multiple statements throughout the report and should be corrected to reflect the FDA-approved labeling for Austedo. These are noted in the detailed comments below.	
10	Teva	The use of a cost-utility model for assessment of value for treatment of tardive dyskinesia is premature, given the state of the scientific evidence for this indication. Further research is needed in this area to develop reliable and valid metrics using a patient-centric approach. ICER’s cost utility approach does not capture the full patient experience. The utilities used were not elicited from a relevant patient population. Further, tardive dyskinesia (TD) is primarily a functional disability, and quality of life literature in this area is scarce; therefore, QALYs are not an appropriate measure for evaluating treatments for TD. A cost-effectiveness approach in the base-case, taking into account reduction in symptoms and improvement in functional measures would be more suitable for this condition. We note that ICER does address this alternative approach, albeit briefly, in the present draft. Until large-scale quality of life studies are conducted among TD patients and functional measures are further developed to capture the extent of disability suffered by TD patients, the cost effectiveness of deutetrabenazine cannot be evaluated	We agree that there is limited data on the impact of TD, or therapies for TD, on patient quality of life. As a result, we chose what we believed to be the best available evidence for the impact of TD on health utilities. As a generous estimate of the potential impact of VMAT2 inhibitors on health utility, we gave the full quality of life benefits of completely eliminating TD symptoms to all patients who had their AIMS score improved by at least 50% in clinical trials. We also conducted extensive sensitivity analyses around the utility estimates, as is recommended in situations where uncertainty exist. The report has been updated to be more transparent regarding how the utility values used in the model were solicited and the limitations of this method. We have also included a comparison of mental health-related and other medical conditions that produce similar disutilities to place these estimates in context. As new evidence emerges, we may develop a "Brief Evidence Update" on an ad hoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other treatments ICER aims to focus its evaluations to inform policy decisions at or near the time of regulatory approval.

	Who?	Comments on TD Draft	ICER Response
		properly. The extant evidence for the TD population is not sufficiently mature to substantiate use of QALYs or a cost-utility framework alone.	
11	Teva	ICER's current approach to calculating direct healthcare utilization costs fails to include the full scope of costs incurred due to having TD. A large study conducted by Ascher-Svanum et al 2008 (737 patients with TD, 1,538 patients without TD) reported that the proportion of patients with paid employment was significantly higher for those without TD versus with TD (23.2% vs 17.7%, p=0.014), and that mean income was significantly higher for patients without TD. Estimates of lost productivity should be applied to patients in the Moderate/Severe TD state (as in scenario analyses [Table H6, page 126].	All base case analyses are conducted from the health care system perspective at the longest feasible time horizon, usually the full lifetime of patients. Including patient productivity costs is not consistent with the health care system perspective of the model, but we do add a scenario to understand productivity gains from a societal perspective, incorporating lost wage estimates from the Ascher-Svanum study. Despite the lack of evidence that having at least 50% reduction in AIMS score allows patients to return to work, we included a benefit of 5.5% of patients (i.e. 23.2% - 17.7%) with improved symptoms being able to reenter the workforce. It is important to note that in the Ascher-Svanum study, patients with TD were paid a higher salary than those without (with TD = $746.70 vs. without TD = $678.60, p=0.016; NS when adjusted for multiple comparisons). Instead of placing those with improved TD at a disadvantage, we chose to apply a generous estimate of the US median salary to the 5.5% of those with improved TD in our scenario analysis.
12	Teva	ICER should also highlight social stigma by including the impact of TD treatments on productivity in the base-case model. Boumans et al 1994 demonstrated that patients with orofacial dyskinesia were less likely to be selected for a job (Boumans et al 1994). Although ICER makes a nod to productivity loss in a scenario analysis losses [Table H6, page 126], given the condition is primarily a functional disability, inclusion of productivity loss in the base-case is warranted.	All base case analyses are conducted from the health care system perspective at the longest feasible time horizon, usually the full lifetime of patients. Including patient productivity costs is not consistent with the health care system perspective of the model, but we do add a scenario to understand productivity gains from a societal perspective. Understanding the potential importance of productivity losses, we chose to include in our report a scenario analysis incorporating lost wage estimates.
13	Teva	These studies highlight the limitations of ICER's value framework for functional disorders. As information on patient-reported	We evaluated current evidence for tardive dyskinesia. However, we acknowledge that evidence is constantly developing and evolving, especially for this condition which has seen a resurgence of research in recent years. We know that reconsideration of

	Who?	Comments on TD Draft	ICER Response
		outcomes develops, new data should be considered in further evaluations. Additional comments on model inputs are detailed below	evidence on comparative clinical effectiveness and value can benefit all stakeholders, including patients, clinicians, payers, and drug manufacturers. Given the availability of new pivotal evidence, we may develop evidence updates on an ad hoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other treatments ICER focuses its evaluations to inform policy decisions at or near the time of regulatory approval.
14	Teva	Use of anticipated real-world compliance and persistence rates with VMAT2 inhibitors in the evaluation: As the majority of patients with TD have underlying psychiatric conditions, the rate of adherence to medications in this patient population is low (Kane et al 2013). This should also be considered in the economic models of any medication use in this patient population.	We agree that the inclusion of long term adherence and persistence is important when estimating long term cost-effectiveness of therapies. However, there is no real world adherence evidence on which to base these estimates. VMAT2 inhibitors treat a symptomatic problem and provide relatively rapid benefits. In addition, they appear to be very well tolerated. They are therefore unlike other medications used in this population, which often have a lag before providing observed benefit and are associated with significant adverse drug events resulting in poorer adherence. Second, poor adherence is often associated with poorer outcomes. It is likely that including adherence in the model would reduce both costs and the VMAT2 inhibitors' effectiveness. Information about how poor adherence to VMAT2 inhibitors affects TD symptoms (i.e. 50% reduction in AIMS or utility) is not currently available. We did include persistence in our base-case model, using estimates from longer-term open-label studies. We also varied these estimates for persistence in sensitivity analyses. It should be noted that higher persistence resulted in better incremental cost-utility. When available, incorporating real-world persistence estimates in the model will likely result in a higher incremental cost-utility ratio.
15	Teva	Appropriateness of including tetrabenazine in the evaluation: Currently, tetrabenazine is not FDA approved for the treatment of TD.	Tetrabenazine has been kept in this report for several reasons. First, this review focuses on VMAT2 inhibitors in TD and tetrabenazine is considered in this class of agents. Second, we received input from clinical experts that tetrabenazine has been used off label to treat patients with TD. These experts felt that tetrabenazine may remain an option for patients with TD symptoms. Finally, though included in this report, the limited data available in the literature did not permit inclusion in models comparing tetrabenazine to other agents or placebo.

	Who?	Comments on TD Draft	ICER Response
16	Teva	[p.8] The report states, “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.” Please note that deutetrabenazine (Austedo) was approved for the treatment of chorea associated with Huntington’s disease. In addition, the FDA-approved labeling for Austedo does not carry the same warnings and precautions as tetrabenazine. This statement should be revised to correct this information. Any reference to the boxed warning throughout this report should include the context of “in patients with Huntington’s disease.”	We have revised this statement.
17	Teva	[p.9] Pricing for deutetrabenazine 24 mg should also be displayed in Table 2.1	Table 2.1 presented the 36mg dose which we used for the economic model. However, we highlighted in the table that the daily dose may range from 12mg to 48mg.
18	Teva	[p. 31]For Table 4.7, the baseline AIMS in mITT population was 9.5 in placebo group, 9.6 in the 12mg group, 9.4 in 24mg group, and 10.1 in the 36mg group; this currently reads "NR	Thank you. We have updated Table 4.7 to show the baseline AIMS score in AIM-TD trial.
19	Teva	[p.32]Table 4.8, ARM-TD, LS Mean AIMS Change from Baseline for deutetrabenazine should have † for the cell value -3.0 (i.e., p value≤0.05); the reported p= 0.019.	Thank you. Table 4.8 has now been revised.

	Who?	Comments on TD Draft	ICER Response
20	Teva	[p.35]The following statement needs to be corrected to accurately reflect the FDA-approved labeling for Austedo for use with patients diagnosed with tardive dyskinesia: "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).	We have revised this statement.
21	Teva	[p.40] The statement, "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…" is incorrect as noted above and should be revised. The boxed warning was not added to Austedo labeling specifically for the TD indication.	We have revised this statement.
22	Teva	[p. 42] The following statement should be revised as it is not consistent with the FDA-approved labeling for patients with tardive dyskinesia, "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."	We have revised this statement.
23	Teva	[p45] ICER should adjust average age of onset and age of treatment to be more specific to US payers; this is especially relevant for a lifetime model. For example, see Loughlin et al 2017	ICER uses a "population" level health system perspective for its base case. The study used for our baseline estimates of age of onset, gender, and underlying condition utilized the Medical Expenditure Panel Survey (MEPS) from 1996 to 2005. MEPS is a randomly selected annual sample of 23,000 - 35,000 non-institutionalized US civilians representing the US population with a low selection bias. Limitations in this data set are that the report included data only up to 2005 (prescribing may have changed) and

	Who?	Comments on TD Draft	ICER Response
			reported antipsychotic use only, not presence of TD (different antipsychotic agents are associated with a different risk of TD). In contrast, the source used by Loughlin et al is an integrated claims and EHR database. Advantages to this data source include being able to identify patients with TD, but not severity of TD. Limitations are that integrated health care systems and employer-based insurance programs are likely over-represented in this data, which is of particular concern given the reliance of many TD patients in the US on public insurance or other assistance programs. We believe that the Domino et al. study (2008) is the most appropriate for the ICER perspective. Given the limitations of the Domino et al. analysis for this model and to identify whether changing patient characteristics had an important impact on the models results, we developed a scenario analysis that incorporated this population.
24	Teva	[p45] ICER's report should reduce the duration of the placebo effect. Currently, patients enter the model based on response rates observed in AIM-TD. Response to treatment remains constant for all responders, and patients do not improve or decline beyond their initial response to therapy while remaining in the "improved TD" state. Patients can only leave the Improved TD state via discontinuation of the study drug. In contrast, in the placebo group, 12% of the population maintains benefit for the lifetime duration, meaning they are guaranteed to never transition to Moderate/Severe TD. The impact of this modeling flaw is exacerbated in scenario analyses, where placebo responders never discontinue their baseline disorder medication, and in turn do not incur added costs or reduced quality of life associated with uncontrolled underlying conditions. The placebo effect should not extend beyond trial	The placebo response is actually a combination of 1) the patient's natural clinical course; and 2) the placebo effect. The treatment effect is a combination of 1) the patient's natural clinical course; 2) the placebo effect; and 3) the treatment effect. To assume that the response observed in a clinical trial is solely due to the placebo effect ignores that the patient condition changes over time. We heard from a number of clinicians that the presence of TD symptoms is dynamic and that the condition often partially resolves in a number of cases. With the lack of published literature describing the natural course of TD, it is not clear whether the natural clinical course of the disease or the placebo effect was most responsible for the observed placebo response. In our model, an appropriate conditional probability was used for patients discontinuing therapy with a VMAT2 inhibitor due to non-persistence, sending some patients who discontinued therapy to a state where they continued to respond (i.e. placebo response). In sensitivity analyses, if all patients discontinued their therapy, the same proportion of patients in the treatment and placebo models received the placebo response, suggesting that there is no flaw in the model conceptualization or design. Furthermore, decreasing the placebo response to 0% in sensitivity analyses had a minimal effect on the increment cost-effectiveness ratio.

	Who?	Comments on TD Draft	ICER Response
		period, as the basis for placebo effect is sense of hope for improvement (Dumitriu, et al., 2010) which should decrease post blinding (Benedetti, et al, 2005).	
25	Teva	[p49] The draft report states that ICER modeled the effects and costs of the highest doses reported in clinical trial (36mg per day) because those doses were "generally associated with the highest effects"; however, with 24 mg dosing, 35% of patients experienced >=50% improvement in AIMS score (Anderson et al 2017), which is greater than 33% improvement observed for the higher dose. The efficacy and corresponding price for the 24 mg dose should then be used in the model for consistency.	With the exception of those with poor CYP2D6 metabolizers, the dosing information for Austedo (deutetrabenazine) recommends titrating the dose up to 48 mg (24 mg twice daily) for TD (Link to: Austedo titration schedule). We therefore believe that the appropriate comparison for the economic model is the 48 mg dose. However, data on the effect of the 48 mg daily dose on 50% improvement in the AIMS score was not evaluated in the Anderson et al (2017) study, as the maximum dose evaluated was 36 mg per day. We have therefore chosen to include the 36 mg daily dose and the corresponding improvement from the clinical trial, as this will provide a very conservative estimate of the incremental cost-effectiveness ratio. We also note that this matches reasonably well with the 38 mg median dose from the ARM-TD titration trial.
26	Teva	[p50] ICER's current approach to calculating healthcare utilization costs is not comprehensive. ICER should consider additional incremental healthcare utilization costs. See, for example, Carroll, et al., 2017 where, post-diagnosis, TD patients had significantly greater annual all-cause healthcare costs versus matched non-TD patients as shown in Figure 3 ($10,199 vs $2,605). Also, any discussion of the cost of TD treatments should appropriately account for real world rates of antipsychotic drug treatment modification or dose reduction and effects on patient quality of life and healthcare costs.	The Carroll et al. analysis was a retrospective cohort analysis of the Truven MarketScan Commercial and Medicare Supplemental Databases presented as a poster at the 2017 Psych Congress in New Orleans. Control patients were propensity score-matched, although the method for matching is not well described. It appears from the demographics table that propensity score matching produced similar groups for age and gender, but not for commercial/Medicare mix or diabetes, which is associated with long-term antipsychotic use. As a result, we believe that the propensity score matching process used in this study was ineffective in balancing known confounders and likely unknown confounders. Since the cost of TD care has been shown in numerous studies to be associated with disease severity, we believe these estimates suffer from confounding by indication and are biased.

	Who?	Comments on TD Draft	ICER Response
27	Teva	[p50] For utility estimation, the state of the science here is not well-advanced, and it would be advised to consider using a CEA, rather than CUA, model for this indication. Alternatively, Briggs et al (2008) utilities for extrapyramidal symptoms may be considered until evidence supporting the perspective of patients with tardive dyskinesia may be further developed.	There is no supporting evidence provided to suggest that utilities for extrapyramidal symptoms reflect those of patients with TD. Also, this study was not submitted as part of a systematic review of the literature. The Briggs et al. study (2008) does support that patients with extrapyramidal symptoms tend to rate their conditions as having a higher utility than do lay persons. Given this evidence, we believe that utilities solicited specifically for moderate to severe TD from a lay population, as was done in Lenert (2004) is the best available evidence for disutility associated with TD. In addition, as a generous estimate of the potential impact of VMAT2 inhibitors on health utility, we gave the full quality of life benefits of completely eliminating TD symptoms to all patients who had their AIMS score improved by 50% in clinical trials. We also conducted extensive sensitivity analyses around the utility estimates, as is recommended in situations where uncertainty exist.
28	Institute for Patient Access	ICER repeatedly attempts to evaluate the cost-effectiveness of a therapy before all the necessary data is available. Such was the case with ICER's draft report on therapies for atopic dermatitis, which were not even priced and publicly available when ICER completed its analysis. Timing is once again a factor in the data available for assessing TD therapies' cost-effectiveness, as detailed in the following pages.	We recognize that for newly approved treatments there is often limited data available. However, patients, clinicians and insurers are still faced with decisions about how best to use these new agents once approved for use. As such, we view comparative clinical effectiveness research, and cost-effectiveness modeling as a useful and important way to identify the key inputs that impact the effectiveness and cost of a new therapy. Even when there is uncertainty about the actual values used in the models, sensitivity analyses can highlight the range of plausible values and their impact on overall cost-effectiveness.
29	Institute for Patient Access	Another recurring concern is whether cost-effectiveness studies and the QALY metric in particular are appropriate and accurate for diseases that are inherently qualitative. A disease such as a cancer, for example, presents finite data points, whether that be the exact size of a tumor or the duration of a patient's remission. Other diseases are not so easily quantified. How does one assign a value to the embarrassment and stigma of, as with	Though we agree that it can be challenging to assign values to the utility underlying the QALY, we believe that this remains the best available way to quantitatively assess the value of a new therapy compared to other active therapies or placebo. In spite of limitations in using QALYs, there remain a number of advantages as well. First, it permits comparing the utility of therapies across conditions. This permits policymakers to assess the relative value of therapies available to patients and clinicians. Second, we are not aware that QALYs have been show to systematically bias against therapies directed at conditions that are not life-threatening vs. those that are. Third, even if it is difficult to assign a value to outcomes that are inherently qualitative, one can assess utilities across a range of plausible values to examine how

	Who?	Comments on TD Draft	ICER Response
		TD, having one's face contort uncontrollably in public? How does one quantify the discomfort of poorly tolerated treatments for psoriasis or the pain and daily inconveniences of rheumatoid arthritis? Treatments for some disease states simply do not lend themselves to economic number crunching.	much the overall cost-effectiveness is dependent on the assigned utility. If the range of utilities do not have a large impact on the value of the treatment, as shown in our models of VMAT2 inhibitors for TD, then one can then focus on those variables that are more important.
30	Institute for Patient Access	Finally, despite ICER's laudable efforts to engage patients and advocacy groups, the framework used to evaluate these patients' therapies has no meaningful way to incorporate their insights. While ICER may relay the patient community's input in its reports, the calculations that result in ICER's benchmark value prices are not designed to quantify patient feedback as a numerical value that impacts the analysis' final findings.	We did seek out input from patients and advocacy groups throughout our review and we believe that our report highlights their insights and concerns. Though it is not possible to include all of these insights into our cost-effectiveness model itself, these quantitative assessments are only one part of our report. We focus considerable attention on the data available, its limitations as well as key insights from all concerned groups including patients and their advocates. Presenting this data, along with insights from patients and other interested parties along with the quantitative results are all necessary to inform policymakers about how best to consider new therapies. The comparative clinical effectiveness, quantitative evaluation, other benefits, and contextual considerations sections of our report all feature prominently in the ICER value framework to inform all decision making by our panels.
31	Institute for Patient Access	Thus, in addition to considering the concerns outlined in the following pages, we urge you also to consider these broader trends and their impact on patient access.	As noted in our prior comment, the results of our cost-effectiveness modeling are only one part of our report. We also have highlighted many of the concerns you and others have raised and believe that these issues are all part of the broader decision about how to best use these new agents for patients with symptomatic TD.
32	Institute for Patient Access	Institute for Patient Access is concerned that ICER's draft evidence report, dated October 2, 2017, undervalues the benefits that tardive dyskinesia (TD) patients can receive from VMAT2 inhibitors. This undervaluation arises because of the reasons described below. 1. The base model does not incorporate the benefit of TD patients' improved adherence to their antipsychotic medicines. As is widely recognized, the physical and psychological	As noted in our prior comments, the base-case model was conducted from the health care perspective. Despite a lack of evidence supporting that improvement, and not elimination, of TD symptoms improves medication adherence, we conducted a scenario analysis further that evaluated the impact of improved adherence to antipsychotic medications. We acknowledge that reconsideration of evidence on comparative clinical effectiveness and value is important for all stakeholders. Given the availability of new pivotal evidence, we may develop an evidence update on an ad hoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other

	Who?	Comments on TD Draft	ICER Response
		impairment caused by TD leads some patients to discontinue their antipsychotic drugs. The draft report acknowledges the costs of poor drug adherence, stating that "sub-optimal adherence or deliberate dose-reduction have been shown to increase the risks of psychotic exacerbation and relapse" (p. 40). Since TD is associated with lower adherence rates to antipsychotic medicines, it logically follows that medicines that control TD could increase patients' adherence to their antipsychotic medicines. The draft report, however, overlooks potential adherence benefits because they have "not been evaluated in clinical studies to date, and so real-world data will be needed to assess these effects." Increased adherence is a fundamental potential benefit of controlling TD. It is inappropriate to assume away this important benefit simply because the novelty of these medicines has provided insufficient opportunities to study the issue. If, as the draft report states, "real-world data will be needed to assess these effects," then ICER should abstain from evaluating the cost-effectiveness of these medicines until such data has been produced. Relegating the important impact of adherence to the scenario analysis, as the report does, is insufficient. Such core issues should be incorporated into the base case results. Further, the scenario analysis employs	treatments ICER focuses its evaluations to inform policy decisions at or near the time of regulatory approval.

	Who?	Comments on TD Draft	ICER Response
		arbitrary assumptions to "account" for non-adherence, so the results cannot be relied upon as a reasonable estimate of the impact that VMAT2 inhibitors will have on patients' adherence to their antipsychotic medicines.	
33	Institute for Patient Access	2.The cost-effectiveness model is biased against VMAT 2 inhibitors. Two issues limit the applicability of the QALY methodology used by the draft report to evaluate VMAT2 inhibitors. First, while improved clinical outcomes are an important benefit of these therapies, so is the enhancement of patients' quality of life. With respect to the draft report, these quality of life benefits are the primary benefit evaluated. However, as documented in a review of the literature that examined the limitations of the QALY methodology, "the QALY system could lead to an innate preference for life saving over life enhancing treatments because preventative or basic long-term care measures generally score lower on QALY calculations than more dramatic treatments. This places certain interventions at a disadvantage – for example those in mental healthcare, where treatment modalities largely fall into the remit of life enhancing measures." Therefore, there is reason to suspect that the QALY methodology underestimates the benefits from VMAT2 inhibitors for patients living with TD.	We acknowledge that a limitation of the QALY is that it may not be fully sensitive to treatment effects, especially where treatment effects are small and not dramatic. However, the beneficial impact of VMAT2 inhibitors on quality of life or utility was not directly assessed. As a generous estimate of the potential impact of VMAT2 inhibitors on health utility, we gave the full utility benefits of completely eliminating TD symptoms to all patients who had their AIMS score improved by 50% in clinical trials. We also conducted extensive sensitivity analyses around these utility estimates. It is likely that with our conservative methods, we have overestimated the utility gains of the VMAT2 inhibitors, rather than underestimated them.

	Who?	Comments on TD Draft	ICER Response
34	Institute for Patient Access	Second, as noted by Hyry et al. (2014), cost-effectiveness assessments are flawed with respect to rare diseases because the small population size, by definition, raises the costs per patient. [1] While TD is not officially a rare disease, its population size (approximately 500,000 patients) is small compared with many other diseases. This size limitation significantly constrains the applicability of the methodology used in the draft report to effectively evaluate the benefits of VMAT2 inhibitors.	It is true that there are likely economies of scale for treating all medical conditions. This is not a flaw of cost-effectiveness analysis, which estimates the actual incremental costs and benefits of therapies, but rather is a reality of the markets. In addition, the population size for TD still does lend itself to adequate sample size and power in clinical trials as well as the use of standard outcome measures. As part of ICER's value framework, we describe possible other benefits and contextual considerations associated with these treatments, and expect there to be robust discussion of these issues at the public meeting.
35	Institute for Patient Access	3. There is an association between tardive dyskinesia and more severe psychopathology. Studies have also found that patients living with TD tend to experience psychological disorders with higher severity than do patients who are not living with TD. For example, in a 2008 study, Ascher-Syanum et al. found that patients with tardive dyskinesia "had significantly more severe psychopathology, were less likely to experience symptom remission, had more severe extrapyramidal side effects, and had lower levels of quality of life and functioning, lower productivity, and fewer activities (all $p < .001$) across the 3-year follow-up. These clinical outcomes impose real costs on patients living with TD that the draft report does not adequately discuss, let alone quantify as a benefit of the medicines that more effectively manage a patient's TD. The value to patients from these medicines that	We agree that there is limited data available on how TD contributes to the cost of caring for these patients and how the use of VMAT2 inhibitors may change these costs. This highlights an area for study as these agents begin to be used in clinical practice. The use of antipsychotic agents includes patients with severe psychiatric disorders as well as a range of other conditions. How the use of VMAT2 inhibitors affects the course of the underlying condition and the associated costs of care remain unknown. Importantly, the timeline of these associations is not known. More severe psychopathology may lead to higher doses of antipsychotic therapy resulting in a higher risk and severity of TD and other extrapyramidal effects. Alternatively, there may be a common pathology or sensitivity for developing TD and extrapyramidal side effects that is related to the psychopathology. However, it is unlikely that the VMAT2 inhibitors will directly affect psychopathology or extrapyramidal side effects. We presented one hypothetical benefit of VMAT2 inhibitors leading to improved control of the underlying conditions in a scenario analysis. Should additional information become available on unanticipated benefits of VMAT2 inhibitors, we may develop a "Brief Evidence Update" on an ad hoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other treatments ICER will continue to focus its evaluations to inform policy decisions at or near the time of regulatory approval.

	Who?	Comments on TD Draft	ICER Response
		treat TD cannot be fully understood without incorporating the potential impact that these medicines can have on improving these clinical outcomes.	
36	Institute for Patient Access	ICER's assumption that there is no association between tardive dyskinesia and increased mortality is likely overstated. As part of the key model assumptions made in the draft report, ICER states that TD does "not have a direct effect on mortality." This may be too strong of an assumption. Chong et al. (2009) examined the mortality rate of 608 Asian patients that were diagnosed with schizophrenia over six years. The study found that while age was a factor, there was "a robust association with increased mortality rate and TD, but we failed to find any significant association with any specific cause of death and TD." A study in a Japanese medical journal back in 1989 also found that schizophrenic inpatients with TD had a significantly higher mortality rate than the inpatients that were not diagnosed with TD. Studies have not universally found a link between TD and a higher mortality rates. However, considering the severity of the outcome, this increased risk potential warrants consideration in the cost-effectiveness assessment when evaluating the potential benefits from new medications for TD – even if there is only a low probability that	We agree completely with the comment that there is an association between TD and early mortality. We have updated the language to better reflect the intent of the statement in the key model assumption section, that there is no evidence supporting improved mortality with these therapies.

	Who?	Comments on TD Draft	ICER Response
		patients living with TD face an increased mortality risk.	
37	Institute for Patient Access	For the above reasons, we have reservations regarding the conclusions of the draft ICER report, and its potentially negative impact on patient access to VMAT2 inhibitors. We encourage ICER to, at a bare minimum, amend the draft report to account for the considerations raised in this letter. Ideally, ICER will reserve judgement on the cost effectiveness of VMAT2 inhibitors until the information deficits identified in these comments are filled with more comprehensive clinical data.	Again, this comment focuses on the results of the cost-effectiveness analyses. The report also highlights the available evidence on the benefits and harms of VMAT2 inhibitors. We conclude that there is promising, but inconclusive evidence supporting the benefits and safety of the two new agents in trials presented to date. We also review other benefits and contextual considerations that are not captured in the quantitative analyses. In terms of the cost-effectiveness of VMAT2 inhibitors, we recognize the limitations of the existing literature used to identify data inputs into the model. However, we also have performed extensive sensitivity and scenario analyses that show the robustness of the cost-effectiveness estimates across a range of plausible inputs.
38	Movement Disorder Policy Coalition	TD is a complex disease characterized by jerky, involuntary movements of the face and body. The loss of physical control in patients with TD can cause those affected to feel embarrassed and may make those around them feel uncomfortable. ICER's draft report goes so far as to classify TD as "extremely debilitating… result[ing] in social isolation." The physical manifestations of TD can lead to compromised mental, emotional and social functioning. The influence of TD on these multi-dimensional domains of being meets the definition of health-related quality of life, an important concept with intangible value, as defined by the federal Office of Disease Prevention and Health Promotion. Unfortunately, there is no known cure for TD. That is precisely why treatments that improve the quality of life for	We agree that TD can have important consequences for patients and that new therapies are needed for this condition. We recognize that the quantitative analyses do not fully capture the qualitative impact of TD on the lives of affected patients. However, we have tried to include a range of plausible inputs in our sensitivity analyses to examine the impact of the input variables on the quantitative model outputs. These results show that our cost-effectiveness ratios are only minimally affected by varying the utility of treating TD across a range of plausible values.

	Who?	Comments on TD Draft	ICER Response
		patients who suffer from this complex disease are significant. As referenced in ICER's draft report, TD is primarily caused by prolonged use of antipsychotic medications and can lead to "decreased compliance with the drugs given to treat the underlying condition." As noted, some patients may try to address their TD by discontinuing their antipsychotics, which can result in ultimately losing control of both conditions. In an ever-evolving and increasingly expensive health technology arena, we all want to obtain the maximum value for health care investments. But ICER's approach struggles to adequately account for the qualitative nature of a disease such as TD. How can we quantify the value of fewer uncomfortable stares, less awkward public encounters and improved social functioning for those afflicted with TD?	
39	Movement Disorder Policy Coalition	The Movement Disorders Policy Coalition respects the need for payers to balance limited dollars with treatment value, but it is critical to consider more than just the bottom line. TD patients and caregivers understand the value of reduced stigma and improved quality of life.	We agree that qualitative and quantitative input is important in considering the value and role of these new agents.

	Who?	Comments on TD Draft	ICER Response
40	Movement Disorder Policy Coalition	The Movement Disorders Policy Coalition released a white paper this month that highlights the impact of movement disorders. It reads, “research has produced innovative drugs in recent years, providing a source of hope and relief to patients and families facing movement disorders.” In addition to the other known health benefits of vesicular monoamine transporter 2 inhibitors, one such source of hope and relief is TD patients’ increased adherence to antipsychotic medicines. The potential benefit of this outcome is great, but time is necessary for data about adherence and effect to be collected and assessed. It would be prudent for ICER to withhold judgement about the cost effectiveness of this treatment until this dynamic can be studied.	We agree with this statement, but also believe that since these medicines are currently available for use by patients, clinicians and payers, reliable information is needed now. This report uses data that is currently available and highlights the limitations of this data as well as the qualitative input of a range of stakeholders.
41	Movement Disorder Policy Coalition	Therapeutic options have historically been limited for patients with TD and other movement disorders. As new options emerge, however, health plan policies that restrict access can make them difficult for patients to obtain. ICER’s findings could be used by health plans to justify their restrictive policies—further impeding patient access to vesicular monoamine transporter 2 inhibitors. A lack of access means a lack of options, interfering with the ability of a doctor and patient to determine and carry out a personalized course of treatment.	The intent of this report is to highlight available data, input from key stakeholders and the value of these new agents as well as the variables that are most important in quantifying these values.

	Who?	Comments on TD Draft	ICER Response
42	NAMI	At the outset, NAMI would like to reiterate concerns regarding the overall scope of this important ICER review. These concerns were initially raised as part of comments NAMI previously submitted on August 16 of this year. First, this ICER review continues to restrict its focus to clinical effectiveness and economic impact. As a result, the review fails to consider the social impacts of living with TD, the impact on family caregivers and how TD contributes to other medical co-morbidities. Again, NAMI remains concerned that the scope of this review excludes consideration of social impacts associated with TD.	We respectfully disagree with this statement. The cost-effectiveness analyses are only one part of this report. It also includes a detailed review of the existing literature as well as input from stakeholders highlighting potential other benefits of these agents as well as key contextual factors that should be considered by policymakers.
43	NAMI	As previously noted, involuntary muscle movements in the mouth and face region, facial grimacing, lip smacking and other symptoms associated with TD carry enormous social stigma. This in turn leads to further social isolation and exacerbation of the negative symptoms associated with psychotic disorders such as schizophrenia. Likewise, in the case of mood disorders such as bipolar disorder the social isolation resulting from TD can further exacerbate symptoms of depression and lack of self-worth. Finally, this social isolation is often associated with sedentary lifestyle, a poor diet and other factors that result in co-morbid chronic medical conditions associated with serious mental illness. The limited scope of this review	We have sought to capture these concerns in our report. We recognize that current data may not fully capture the experience of patient's living with TD. We have highlighted these concerns in the report itself and have performed sensitivity analyses that include a range of values that may capture these hard to quantify aspects of TD.

	Who?	Comments on TD Draft	ICER Response
		means that the findings of this review fail to capture this patient experience for individuals living with TD.	
44	NAMI	These findings raise the serious prospect that individuals experiencing moderate to severe TD that are diagnosed with schizophrenia, schizoaffective disorder and bipolar disorder will face enormous challenges accessing both Valbenazine and Deutetrabenazine. NAMI is extremely concerned that the very high incremental cost utility, lifetime horizon for both Valbenazine and Deutetrabenazine as compared to the same for placebo will result in both public and private payors refusing to offer access to these promising therapies for which there is no FDA approved alternative. As noted in previous comments to ICER, these new therapies to treat TD amount to a "game changer" for patients that have been living with this condition for years. To see a promising therapy taken from them after many years will be extremely frustrating.	This comment focuses on the results of the cost-effectiveness analyses. The report also highlights the available evidence on the benefits and harms of VMAT2 inhibitors. We conclude that there is promising but inconclusive evidence supporting the benefits and safety of the two new agents in trials presented to date. We also review other benefits and contextual considerations that are not captured in the quantitative analyses. In terms of the cost-effectiveness of VMAT2 inhibitors, we recognize the limitations of the existing literature used to identify data inputs into the model. However, we also have performed extensive sensitivity and scenario analyses that show the robustness of the cost-effectiveness estimates across a range of plausible inputs.

	Who?	Comments on TD Draft	ICER Response
45	NAMI	In describing "stopping/changing" of antipsychotic treatment, this ICER review ignores the risk associated with such an approach and the lack of evidence supporting its effectiveness. In real life clinical practice, reducing, replacing, or removing antipsychotic treatment can jeopardize recovery and stability for people living with a serious mental illness. Interrupting treatment also increases the risk of an episode of acute psychosis, mania and suicidal ideation. Further, it fails to recognize the importance of choice and autonomy for individuals and the value of shared decision-making between prescriber and patient. It is troubling that ICER continues to integrate this into the review as an acceptable treatment option for TD patients.	This report does not advocate for or include such approaches in our analyses. We only discuss this in the setting of comments from clinical experts and existing specialty guidelines about the management of TD.

	Who?	Comments on TD Draft	ICER Response
46	NAMI	The measures used in the final report are based almost exclusively on blinded video recorded expert central scoring. The AIMS (Abnormal Involuntary Movement Scale) is a widely accepted measure for assessing symptoms and therapeutic improvement for patients living with TD. It is important that this review measured AIMS at baseline and follow-up. Unfortunately, the review then used a substantially less reliable process for assessing outcomes – a blinded "expert" review of videotaped interactions between clinicians and their patients with central scoring. In NAMI's view, this raises questions about whether or not the "experts" reviewing videotaped interviews are assessing the performance of the clinicians in accurately diagnosing and prescribing treatment or, instead, the symptoms and outcomes for the patients themselves. NAMI is always concerned when studies rely exclusively on clinician-reported outcomes. While many clinicians may have expertise in diagnosis and treatment, they too often have only brief interactions with their patients (in psychiatry, a "medication assessment" can be as brief 10 to 15 minutes). In this case, it appears that the principal outcome measure is based on review of videotaped interviews.	We present the range of outcomes used in the trials of the new VMAT2 inhibitors. The AIMS was used as the primary outcome in all trials. However, other outcomes included patient and clinical global impression. Though we agree that patient reported outcomes for a condition such as this are very important, the magnitude of benefit observed with patient reported outcomes was in fact smaller than that provided by the AIMS.

	Who?	Comments on TD Draft	ICER Response
47	NAMI	It appears that this review never afforded the opportunity for the "experts" convened by ICER to talk directly with clinicians prescribing treatments for TD in order to assess their opinion about the value of a breakthrough treatment to help their patients. Why was no weight given in the review to the judgment of a clinician who is finally able to prescribe a disease-modifying therapy for a patient that has lived with TD for years?	We did speak with leading clinical experts involved in the care of patients with TD. They did highlight the need for new therapies and the potential impact that such new therapies could have. At the public meeting, there will be a range of stakeholders invited including those mentioned here.
48	NAMI	There is a nearly complete absence of patient-reported outcomes or attempts to measure the patient experience of living with TD. For NAMI, this is the most serious flaw in the ICER review; namely, the complete absence of any patient-reported outcomes. NAMI understands that in the latter stages of this review ICER undertook a survey instrument for people living with TD and their family members. NAMI was grateful for the opportunity to provide input on this survey instrument. Unfortunately, this patient survey came after ICER had already designed and executed this review, rather than seeking input from patients and their families upfront. As noted above, this review could have benefited from upfront input integrating the direct experience of people living with a moderate to severe facial tick or an involuntary movement disorder.	As previously mentioned, this report presents evidence from the existing literature. We captured all relevant outcomes we could including those from patients. We also appreciate the collaboration with NAMI in circulating the survey among their members living with TD and their family. The timeframe of this survey came after seeking input from a range of stakeholders in developing the scope of the review and then in developing the protocol for the evidence review and modeling. Thus, the patient survey followed this initial work. We do recognize that the time frame ICER uses in creating our reports may limit the extent of primary data collection that can be done and appreciate NAMI's input and role in developing and fielding this survey.

	Who?	Comments on TD Draft	ICER Response
49	NAMI	NAMI is hopeful that the planned December 5 meeting will include presentations from people living with TD, as well as opportunities for response to formal presentations of findings from people living with TD (or in the alternative, the perspective of family caregivers). It is critical that this voice be part of ICER's final deliberations.	Our meetings are public and are open to all patients who register in advance to deliver public comments. We also invite individual patients and/or patient advocates to participate as experts throughout the day to advise our voting panel.
50	NAMI	QALYs as a major outcome measure has significant limitations in capturing the experience of living with TD. NAMI recognizes that ICER has traditionally relied on QALYs as a critical measure in assessing value, effectiveness and utility when comparing competing clinical interventions. However, in the case of TD, the use of QALYs significantly fails to capture the complexities of the patient experience. The final results accurately note that the risk of mortality associated directly with TD is rare. We know that an adult diagnosed with moderate to severe TD can live the disorder for decades. Further, being able to effectively control symptoms to the point of maintaining successful employment, peer and family relationships and other aspects of community integration are of high value to patients. Unfortunately, QALYs are largely ineffective in capturing these high value goals to patients. As a result, the ICER fails to effectively capture this patient experience. At the same, the innovative therapies to treat TD that are in the ICER	We agree that there is limited data on the impact of TD, or therapies for TD, on the aspects of patient quality of life outlined in this comment. We chose what we believed to be the best available evidence for the impact of TD on health utilities. As a generous estimate of the potential impact of VMAT2 inhibitors on health utility, we gave the full quality of life benefits of completely eliminating TD symptoms to all patients who had their AIMS score improved by 50% in clinical trials. We also conducted extensive sensitivity analyses around the utility estimates, as is recommended in situations where uncertainty exist. The report has been updated to be more transparent regarding the how the utility values used in the model were solicited and the limitations of this method. We have also included a comparison of mental health-related and other medical conditions that produce similar disutilities to place these estimates in context. As new evidence emerges, we may develop an evidence update on an adhoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other treatments ICER focuses its evaluations to inform policy decisions at or near the time of regulatory approval.

	Who?	Comments on TD Draft	ICER Response
		review are, in NAMI's view, penalized severely in this review for precisely the same reason – that many patients are faced with the prospect of relying on these therapies not as a curative intervention, but as a way of effectively managing their symptoms to promote recovery and integration. With ICER now finalizing and publishing these findings, NAMI would like to express our strong disappointment that this report is likely to be used by payors – both public and private – to block access to innovative therapies that people living with TD have been waiting for decades.	
51	Neurocrine	Topic – TD disutility assumption – (Page 57)- The report states that "a utility decrement of 0.095" was applied "to those patients with moderate to severe TD." As we and others have stated this utility decrement woefully underrepresents the burden of TD and thus undervalues a safe and effective treatment (e.g., valbenazine). Attempting to assign a disutility value that was derived from a study that did not concurrently assess TD severity is biased and underestimates the impact of TD. The Lenert et al. (2004) study did not provide an appropriate estimate of the disutility associated with TD as it was obtained from a sample of the general population. As noted on pg. 6 of ICER's report, the impact of TD is	We agree that there is limited data on the impact of TD, or therapies for TD, on the aspects of patient quality of life outlined in this comment. We chose what we believed to be the best available evidence for the impact of TD on health utilities. As a generous estimate of the potential impact of VMAT2 inhibitors on health utility, we gave the full quality of life benefits of completely eliminating TD symptoms to all patients who had their AIMS score improved by 50% in clinical trials. We also conducted extensive sensitivity analyses around the utility estimates, as is recommended in situations where uncertainty exist. The report has been updated to be more transparent regarding the how the utility values used in the model were solicited and the limitations of this method. We have also included a comparison of mental health-related and other medical conditions that produce similar disutilities to place these estimates in context. As new evidence emerges, we may develop an evidence update on an adhoc basis. Still, given the critical early decisions that are made regarding pricing, coverage, and use of new technologies, in its reports on tests, drugs, and other treatments ICER focuses its evaluations to inform policy decisions at or near the time of regulatory approval.

	Who?	Comments on TD Draft	ICER Response
		not likely to be correctly assessed by people who are not affected. To offer a credible alternative, we are analyzing data from the Medical Expenditure Panel Survey to estimate the impact of TD (ICD9 code 333.xx) on patients' health-related quality of life. Our preliminary results show that the mean EQ-5D index score (utility) for respondents with TD is 0.625 and the mean utility for propensity-score matched respondents without TD is 0.750 (a difference of 0.125). The EQ-5D data was collected in 2000–2003. In addition to EQ-5D utilities, we also examined the SF-12 scores, which were collected from 2000–2015. Respondents with TD scored lower on both the Physical Component (38.4 vs. 41.8) and Mental Component (47.3 vs. 48.1) Scores. We plan to convert the SF-12 scores to utilities using published algorithms. This will allow us to present the TD utility decrement based on a larger sample and for a greater number of years. We will be submitting these analyses for upcoming conferences (AMCP, ISPOR). We argue that the TD disutility of 0.095 used in the current model is too small to correctly capture the impact of TD. We suggest the utility decrement of at least double the value used is more appropriate, as supported by our recent analyses.	

	Who?	Comments on TD Draft	ICER Response
52	Neurocrine	Topic: Proportion of Responders to Placebo (Pages 48–49) The proportion of responders to placebo is a key parameter determining the distribution of patients in the discontinued-improved and discontinued-not improved health states over time. Although the incremental cost per QALY (vs. placebo) results for valbenazine and deutetrabenazine were not directly compared (as noted on of page 19 of the report), we would still like to point out the difference between the placebo values for valbenazine (8.7%) and deutetrabenazine (12.0%). The treatment-placebo differences for valbenazine and deutetrabenazine are 31.3% and 21.1%. This should be captured as a differentiating data point favoring valbenazine. Although a larger placebo effect should work against an inferior product, in this model it seems to work to its advantage via improved TD without treatment/costs. Within the current draft of the model, every year patients discontinue treatment (8.7% of the valbenazine and 12.0% of the deutetrabenazine) yet maintain their improved status. This methodology inappropriately disadvantages valbenazine. We request that you revise the methodology so that placebo transitions from "response no	We used data as reported directly from clinical trials. Differences in the placebo responses likely resulted from differences in the underlying patient populations (e.g. severity of TD, duration of TD, use of other therapies, dose of neuroleptics, adherence to therapy or placebo) and evaluation of outcomes. As such we chose not to indirectly compare the two drugs with each other. The same caution should be applied when comparing the point estimates obtained for the incremental cost-effectiveness ratios. In order to evaluate whether our methodology inappropriately disadvantaged valbenazine, we conducted additional analyses. In the base-case, an appropriate conditional probability was used in the model for patients discontinuing therapy with a VMAT2 inhibitor due to non-adherence, sending some patients who discontinued therapy to a state where they continued to respond (i.e. placebo response). The placebo rate used was the reported rate from the respective clinical trials. Changing this placebo rate resulted in small changes to the incremental cost-utility ratio, demonstrating that valbenazine was not disadvantaged by the methods of our analysis.

	Who?	Comments on TD Draft	ICER Response
		treatment" to "moderate TD" (comparable in both medications).	

	Who?	Comments on TD Draft	ICER Response
53	Neurocrine	Topic: Discontinuation rates-(Pages 48–49; Page 48, Table 5.2; Page 54 Table 5.6) The first-year discontinuation risk for valbenazine (19.0%) "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." Unfortunately, the authors have combined "risks" and "rates" and, as a result, incorrectly computed the annual discontinuation risk, which should be 18.3% for valbenazine. It appears that the authors annualized the difference between the 6-week risk reported by Remington (6.0%) and the 6-month risk reported by Hauser (13.6%). This method is incorrect, because risks occurring over different time periods cannot be combined (subtracted, in this case). Instead, the authors should have converted the risks to rates and then calculated the annual discontinuation risk from the rate difference: • Rate computed from Remington: 0.024364. • Rate computed from Hauser: 0.04125. • Rate difference: 0.016887. • Annualized discontinuation risk: 18.3%.	The comment is correct in that in annualizing probabilities, they should be converted to a rate and the new probability calculated at the new time point. We have corrected the discontinuation rate for valbenazine using the following approach. We first calculated the probability of discontinuing therapy in longer time periods by taking the discontinuation probability at 6 months (the longest period available) and subtracted out the proportion who discontinued therapy at 6 weeks. This resulted in a probability over a 20 week period (from week 6 to month 6) of 7.17% (13.5% discontinued therapy by month 6 - 6.3% who had already stopped therapy by week 6). Those 7.17% of people are those who were still taking the therapy at week 6, but subsequently stopped taking the therapy by month 6, a period of 20 weeks. The resulting discontinuation probability at one year, adjusted using the probability to rate to probability method, is 17.59%. We have adjusted the model accordingly.

	Who?	Comments on TD Draft	ICER Response
54	Neurocrine	The first-year discontinuation risk for deutetrabenazine (13.0%) "was equal to the exposure-adjusted incidence rate reported in Anderson et al. (2017), which only included discontinuations after a "washout period." The Anderson et al. (2017) abstract notes that the exposure-adjusted incidence rate (per patient-years) for discontinuations was 18 per 212.4 person years (»0.085 per person-year). However, there is insufficient information given in the abstract and in the report for us to replicate or verify these calculations. However, as noted above, the differences in discontinuation risks, which were due to adverse events, may reflect differences between the participants of those trials. One key difference in the study populations was the exclusion of anticholinergics in the deutetrabenazine trials. This significantly changes the "risk" for adverse events/discontinuation.	Due to differences in the study populations and clinical trial designs, we chose not to directly compare deutetrabenazine with valbenazine. We used data from the best available reports of the clinical trials and extension studies. Given that data was often available in poster format only, we used sensitivity analyses to examine the effects of uncertainty in the models. Under all sensitivity analyses and scenarios, both VMAT2 inhibitors displayed incremental cost-effectiveness ratios well above commonly accepted thresholds.
55	Neurocrine	Please provide a justification for using such high first-year discontinuation probabilities, which are higher than those used for the preliminary results presented in August. Please also provide a justification for assuming that the probabilities are 50% smaller during subsequent years, rather than some larger proportion. We ask that you consider reducing the first-year Valbenazine discontinuation risk as well as reducing the	The first year discontinuation probabilities were annualized from reports of discontinuation rates from the extension studies, which on average were less than one year in duration. Additionally, patients tend to discontinue therapy at a lower rate in clinical trials than they do in real world settings. We arbitrarily assumed a 50% reduction in the discontinuation rate based on a broad knowledge that discontinuation rates decline over time. A 75% relative reduction in the discontinuation rate is equally arbitrary. To determine the impact of our assumption on the incremental cost-utility ratio, we varied the relative reduction from 0%-95% for the final report.

	Who?	Comments on TD Draft	ICER Response
		discontinuation risk in subsequent years by 75%.	
56	Neurocrine	According to our modeling efforts, we would expect the following to be reasonable scenarios: • 2x disutility-could result in reduction of Valbenazine cost/QALY~ 50% • Placebo transitions from response no treatment to moderate TD (equal to Valbenazine)- could reduce Valbenazine cost/QALY ~20% • Proposed discontinuation in first year and 75% reduction in subsequent years-could reduce Valbenazine cost/QALY ~6% **Combination of the above:** 2x disutility, and Placebo transitions to moderate TD at 5% annually, 75% reduction in discontinuation in years 2+	We used what we believed to be the best available evidence and assumptions for the base case. All of these factors were evaluated extensively in one-way sensitivity analyses.
57	Neurocrine	This intervention provides significant direct patient health benefits that are not adequately captured by the QALY. Correct. The body regions most impacted by TD symptoms are the face/mouth/jaw, limbs and trunk. Recent surveys as well as interim data from our RE-Kinect study show that some of the most bothersome sequalae of tardive dyskinesia (TD) include inability to eat, difficulty breathing, and difficulty in movement/walking. ICER's current utility degradation is based off non-TD sufferers reporting and absolutely does not account for	As stated in the report, we agree that there maybe additional benefits to the patient that are not sufficiently captured by QALY. However, we have tried to include a range of plausible inputs in our sensitivity analyses to examine the impact of the input variables on the quantitative model outputs. These results show that our cost-effectiveness ratios are only minimally affected by varying the utility of treating TD across a range of plausible values.

	Who?	Comments on TD Draft	ICER Response
		the physical impact of symptoms. The report does recognize the potential impact on socialization and ability to work or go to school but underrepresents the impact on the patient and/or caregiver.	
58	Neurocrine	This intervention offers reduced complexity that will significantly improve patient outcomes. Correct. Valbenazine is effective and well tolerated with a single daily dose. Additionally, modeling shows there may be a synergistic relationship with antipsychotics that may allow for optimization of the antipsychotic treatment. With regards to outcomes, we believe valbenazine can (and already has) significantly improve outcomes in many patients. Given ICER's recognition of the importance of patient-centric outcomes, the individual health state will influence the impact of treatment. As we have previously stated, highly functioning patients with limited comorbidities should benefit more than individuals with limited capacity and high comorbidity burdens. We will analyze data from our RE-Kinect study early next year to explore this further.	Thank you for your comments. We highlighted the once daily dosing of valbenazine in our report, which may potentially improve adherence to treatment. However, there is currently no published data showing how the use of valbenazine affects the course of the underlying condition. Similarly, published data to date does not currently identify subgroups of patients who are more likely to benefit from VMAT2 inhibitors. Though it is possible that certain patient groups, such as highly functioning individuals with limited co-morbidities may benefit more from use of VMAT2 inhibitors, this remains to be demonstrated.

	Who?	Comments on TD Draft	ICER Response
59	Neurocrine	This intervention will reduce important health disparities across racial, ethnic, gender, socio-economic, or regional categories. Possibly. Our data analyses, and those of other groups, show that the main risk for the "average" TD patient is determined by cumulative antipsychotic exposure rather than ethnicity or gender. The hope is that our educational efforts will result in renewed interest in the screening and diagnosis of TD while reducing the human and societal burden of TD. Additionally, we hope that the body of evidence for valbenazine and evaluations like the ICER report will allow for similar access to an effective and well tolerated treatment for TD.	Thank you for your comment. We agree with you that there are currently no data showing health disparity across race, ethnicity, gender, socio-economic or regional factors in patients with TD.
60	Neurocrine	This intervention will significantly reduce caregiver or broader family burden. We believe it can. Currently there are few data available regarding caregiver burden. One of the key cohorts in our RE-Kinect study is the patient caregiver. In the study, we gather data on the burden of providing care for the patient with TD. We will analyze and present this data early next year.	Thank you for your comment. As stated in the report, we agree that the use of VMAT2 inhibitors may decrease caregiver/family burden, particularly for individuals with disabling TD. However, due to the lack of data at this time, it is difficult for us to quantify the extent of this impact. Instead, we described this as a potential benefit of the medication. At our meeting, we will discuss in depth other benefits and contextual considerations that we do not currently have any evidence for and will consider this as part of our evaluation of the value of the drug. We look forward to reviewing the RE-Kinect data when available to help us to quantify the impact of this drug on caregiver/family burden.

	Who?	Comments on TD Draft	ICER Response
61	Neurocrine	This intervention offers a novel mechanism of action or approach that will allow successful treatment of many patients who have failed other available treatments. Correct. The term "failed" can imply either intolerance or lack of efficacy (or both). The ICER draft evidence report acknowledges the lack of evidence for available "off label" treatments. The mechanism of action for valbenazine is novel and allows for single daily dosing as well as a favorable NNT/NNH profile. Other VMAT2 inhibitors are tethered to a pharmacologic risk profile (e.g. black box warning and contraindications) due to other deleterious active metabolites.	Thank you for your comment. As stated in the report, we believe that the approval of these drugs for tardive dyskinesia represents a potentially important advancement for individuals with TD.
62	Neurocrine	This intervention will have a significant impact on improving return to work and/or overall productivity. Possibly, if the TD symptoms resulted in either a physical limitation or self-isolation of the patient. These are a few of the many deleterious outcomes highlighted in recent TD patient/caregiver surveys. Some of our clinical trial patients have commented on the reduction of symptoms "allowing me to return to school/work". There will be a case study (by Josiassen) published soon to support this. Subsequent real-world studies (retrospective and prospective) will assist in confirming this benefit.	We agree that it is possible for this intervention to have a significant impact on improving return to work and/or overall productivity. However, current studies of VMAT2 inhibitors have not evaluated return to work or productivity outcomes. At our meeting, we will discuss in depth other benefits and contextual considerations that we do not currently have any evidence for and will consider this as part of our evaluation of the value of the drug.

	Who?	Comments on TD Draft	ICER Response
63	Neurocrine	Other important benefits or disadvantages that should have an important role in judgments of the value of this intervention. As stated previously, there are several components that should be considered to determine “value”: 1) the location and severity of the symptoms being treated; 2) the social impact of the location and severity of symptoms; 3) the physical impact of the location and severity of symptoms; 4) the societal impact of the location and severity of symptoms; 5) the relative healthcare burden of the location and severity of symptoms; 6) the caregiver burden; 7) the impact of the treatment on all the above; and 8) the impact of the treatment on any comorbidities, especially the underlying mental illness.	At our meeting, we will discuss in depth the comparative clinical effectiveness, other benefits, and contextual considerations which feature prominently in our evaluation and in our value framework.
64	Neurocrine	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. Yes (with regards to QOL and possibly length of life if the symptoms manifest in ADL sequelae). For all the reasons stated previously, the physical and social impact of the symptoms can be very significant.	As noted in our report, TD can be a chronic, disabling condition. If future evidence shows that VMAT2 inhibitors provide long-term benefit for patients with persistent TD symptoms, this will favorably affect patient's quality of life over many years.

	Who?	Comments on TD Draft	ICER Response
65	Neurocrine	This intervention is intended for the care of individuals with a condition that represents a particularly high lifetime burden of illness. Yes. The symptoms are brought about by prolonged exposure to antipsychotics and tend to be permanent if not treated. The condition becomes chronic in most patients with TD lasting for decades. The risk of lifetime burden of illness is increasing because of the use of these agents in younger patients with conditions other than schizophrenia.	As noted in our report, continued use of antipsychotic medications despite the presence of TD symptoms may be appropriate for certain psychiatric conditions. For patients with TD symptoms of recent duration and for indications other than schizophrenia, alternatives to the use of antipsychotic drugs may be considered.
66	Neurocrine	This intervention is the first to offer any improvement for patients with this condition. Correct. Valbenazine is the first and most effective treatment approved for patients with TD.	We agree. We absolutely recognize that valbenazine is the first to be approved for this indication by the FDA-- and may offer promising benefits for patients living with TD.
67	Neurocrine	Compared to surveillance with no maintenance therapy, there is significant uncertainty about the long-term risk of serious side effects of this intervention. Incorrect. Our development program has assessed long term efficacy and safety. We have completed 3 long term extension studies with valbenazine for 1 year or longer. The Kinect 3 LTE, Kinect 4 and 1506 roll-over studies all evaluated durability of effect and safety for a year or more. It is estimated there are >350 patient years of moderately long-term exposure from those three trials (i.e., subjects treated for > 3 months up to 2 years). Given the estimated TD patient population of 100-300k within the	Thank you for your comment. First, we want to note that this statement has been revised to say "compared to usual care, there is significant uncertainty about the long-term risk of serious side effects of this intervention". We reported on the 1 year data on efficacy and safety studies of valbenazine as part of our evidence review. However, these studies are single arm studies. The only comparative data available for valbenazine is to placebo, and the RCT phase was for 6 weeks. Therefore, we remain uncertain about the significant long-term risk of serious side effects or magnitude or durability of the long-term benefits, compared to usual care (including use of off-label medications).

	Who?	Comments on TD Draft	ICER Response
		ICER report, this is a substantial accumulation of data with no additional safety signals emerging.	
68	Neurocrine	Compared to surveillance with no maintenance therapy, there is significant uncertainty about the magnitude or durability of the long-term benefits of this intervention. Incorrect. As stated in the previous question. The valbenazine development program has assessed long term efficacy and safety. The Kinect 3 LTE, Kinect 4 and 1506 roll-over study all evaluated durability of effect and safety for a year or more. It is estimated there are >350 patient years of moderately long-term exposure from those three trials (i.e. subjects treated for > 3 months up to 2 years). The data supports persistent and durable effectiveness.	See comment above. Please note that this statement has been revised to say "compared to usual care, there is significant uncertainty about the magnitude or durability of the long-term benefits of this intervention".
69	Neurocrine	There are additional contextual considerations that should have an important role in judgments of the value of this intervention. Agree. As stated previously, there are several components that should be considered to determine “value”: 1) the location and severity of the symptoms being treated; 2) the social impact of the location and severity of symptoms; 3) the physical impact of the location and severity of symptoms; 4) the societal impact of the location and severity of symptoms; 5) the relative healthcare burden of the location and severity of symptoms; 6) the caregiver burden; 7) the impact of the	At our meeting, we will discuss in depth the comparative clinical effectiveness, other benefits, and contextual considerations which feature prominently in our evaluation and in our value framework.

	Who?	Comments on TD Draft	ICER Response
		treatment on all the above; and 8) the impact of the treatment on any comorbidities especially the underlying mental illness.	
70	Neurocrine	We appreciate your consideration of our suggestions. We believe that valbenazine is a significant advance in treating a terrible drug induced movement disorder - tardive dyskinesia. We have been transparent regarding our upcoming data sets that will support our suggestions. Unfortunately, they will emerge weeks or months after the final report. We expect that ICER will include these data sets and revise their analysis to more accurately acknowledge the benefit of treating TD with Valbenazine.	We appreciate the effort of all the organizations that provided input into this review process and for their comments on the draft report. We look forward to presenting the final report at our public meeting and engaging in a discussion with stakeholders about our results as well as other potential benefits and contextual factors.