Project Title: First and Second Generation Antipsychotics in Children and Young Adults-Comparative Effectiveness Review Update

I.   Introduction

This is an update of a comparative effectiveness review (CER) of first and second generation antipsychotics in children and young adults 24 years of age or younger. The original report, published in 2012,[1] underwent surveillance by the Agency of Healthcare Research and Quality (AHRQ)[2] finding new studies and interventions which were anticipated to change several of the original conclusions.

There is no consensus on the terminology to describe antipsychotic medications (e.g., first and second generation, typical and atypical). For the purposes of this review, the terms first generation (FGA) and second generation (SGA) antipsychotics will be used.

Changes Since the Original Comparative Effectiveness Review

The scope of the review has remained quite similar, with key changes being the addition of (1) three newly approved SGAs and a previously discontinued FGA, (2) some conditions of interest, and (3) the assessment of harms by condition (previously only across conditions). The Key Questions (KQ) from the original CER were reviewed by a stakeholder panel and underwent a public comment process via the AHRQ Effective Health Care Program website. There have been a few changes to the KQs. Rather than distinguishing between benefit outcomes primarily by type of outcome (symptom vs. other outcomes), they will be reported by timing and importance to patients; there is now only one KQ for benefits. Moreover, to enhance reporting on subgroups the previous KQ on subgroups has been integrated into the KQs on benefits and harms. Because harms are going to be addressed both by condition and across all conditions, this is made explicit with two different KQs. The original CER used terminology specific to the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV), and the conditions for this update have been revised according to changes in the DSM-V (e.g., pervasive developmental disorders is currently classified as an autism spectrum disorder) published in 2013.[3] None of these changes are anticipated to impact the categorization or inclusion of previous studies for this update. Diagnosis of study participants based on DSM-V will not be mandatory for the update CER. Specific changes are described below in terms of the PICOTS (population, intervention, comparators, outcomes, timing, and setting).

Population. Apart from nomenclature, there has also been the (1) addition of depressive disorders, anxiety disorders, and substance use disorders; (2) broadening of anorexia nervosa to include other eating disorders, and of Tourette's syndrome to include all tic disorders; and (3) specification that the category of behavioral issues includes treatment of symptoms outside the context of a disorder, as for example when antipsychotics are prescribed for sedation/sleep within certain environmental contexts (e.g., residential facilities). While these latter uses of antipsychotics are not endorsed by guidelines or indicated for antipsychotic use as per FDA approval, it was thought important by our stakeholders to review the evidence on all current uses of antipsychotics to provide information of benefit and harms for a broad range of stakeholders.

The subgroups have been modified slightly to include phase and features of disorder (e.g., acute mania vs. maintenance treatment in bipolar disorder, first-episode psychosis vs. treatment in context of prior episodes, presence of psychosis in disorders apart from schizophrenia), medication dose, and use for cases of refractory treatment; these reflect some major components of the uncertainty currently faced by many clinicians. We have indicated the difference between patient- and intervention-level characteristics (i.e., dose and co-interventions).

Interventions and Comparators. One long-standing FDA-approved FGA (molindone) was discontinued at the time of the original CER, but a generic has recently received approval for marketing and therefore this FGA has been added as an eligible antipsychotic. The SGA lurasidone was approved by the FDA in 2010 (for schizophrenia and later for bipolar depression, both in adults) and was not reviewed in the original CER. Two other SGAs were approved in 2015: brexpiprazole in July for schizophrenia and adjunctive treatment of major depression in adults, and cariprazine in September for schizophrenia and bipolar disorder in adults. The comparators remain the same: placebo/no treatment, same antipsychotic of different dose, and another antipsychotic.

Outcomes. There have been changes to the terminology and classification of some outcomes, for example removal of the wording "patient- or family-reported outcomes" from a single outcome, because several of the outcomes are measured by patient/family report. Despite changes, all of the previous included outcomes will be captured in some manner. There has been the addition of an outcome for global impressions, which captures symptoms and overall clinical improvement. The outcomes related to harms have been modified slightly to have better consistency with the categories of major and general adverse effects. The outcomes that will be graded for strength of evidence have been modified to be more precise for symptoms that are treated with antipsychotics for each condition (e.g., "autistic symptoms" has been replaced with irritability) and to reflect any changes to terminology and classification.

Timing and Setting. The same criteria will be used for timing (1987 or later) and setting (all settings). Outcomes will be categorized in terms of short- (<6 months) and long- (≥ 6 months- <12 months; 12 months+) term followup.

II.   Background, Rationale, and Objective for this Comparative Effectiveness Review Update

The use of antipsychotics in children, youth, and young adults has risen dramatically over the past 20 years,[4-9] with the prescribing frequency in the United States increasing from 8.6 per 1000 children in 1996 to 39.4 per 1000 in 2002.[10] Annual sales of the newer class ("second generation") of antipsychotics (see below) in 2010 were $16.1 billion, growing by $1.4 billion since the previous year.[11] This drug class had also become the most costly within the Medicaid program, far exceeding the costs of any other drug class.[12]

Antipsychotic medications are commonly categorized into two classes. FGAs were developed in the 1950s, while SGAs emerged in the 1980s. Each class is considered to have a distinct side-effect profile, although there is considerable overlap between them. FGAs are mainly associated with dry mouth, sedation, and extrapyramidal symptoms, which are movement disorders characterized by repetitive, involuntary muscle movements, restlessness, or an inability to initiate movement. Neuroleptic malignant syndrome is a rare but serious adverse effect. In the United States there has been a near disappearance of the use of FGAs over the last two decades.[13]

A shift towards the use of SGAs was partly driven by their lower risk of extrapyramidal symptoms and other adverse events caused by the persistent dopamine receptor blockade by FGAs. The pharmacology of SGAs is diverse (based on action at several types of receptors) with associated heterogeneity in effects and harms; nevertheless, this class is more prone than FGAs to some adverse effects including weight gain, elevated lipid and prolactin levels, and development of metabolic syndrome.[14-16] This risk profile has led to great concern, because of the known associations between weight gain and obesity with diabetes, dyslipidemia, and hypertension, all of which are leading risk factors for future cardiovascular morbidity and mortality.[17] Together with a longer duration of use of SGAs than FGAs, this risk profile necessitates safety monitoring and prescription choices based on benefit-risk assessments.

For most FGAs and SGAs, the U.S. Food and Drug Administration (FDA)–approved indications for children (≤ 18 years of age) are restricted to the treatment of schizophrenia and bipolar mania. Other pediatric indications approved by the FDA include treatment of irritability associated with autism in children 5 years or older (risperidone in 2006 and aripiprazole in 2009) and of Tourette's syndrome in children aged 6-18 (aripiprazole in 2014) or over 8 years (pimozide). Off-label prescriptions are common in children and adults.[18] Approximately 31 percent of antipsychotic treated youth have ADHD,[19] and 34.5 percent of antipsychotic treated young adults have depression.[9] In Medicaid-enrolled children, ADHD accounted for 50 percent of total antipsychotic use in 2007, and ADHD together with mood disorders not otherwise specified were the most common uses (32 and 37.2 percent, respectively) for antipsychotics in a sample of Medicaid-insured children in Vermont during 2012.[18] In these cases or other conditions such as conduct disorders or depression, antipsychotics are usually given for adjunctive treatment of severe behavioral symptoms (e.g., aggression), rather than for psychoses.[9, 12] They may also be prescribed for mood instability or relatively minor symptomatology (e.g., insomnia) of a condition, or even outside the context of a condition;[18] these uses are accompanied by considerable controversy because of concerns regarding the balance of benefits and harms.

Because of the marked increase in on-label and off-label use of antipsychotics, prescribing practices have been under ongoing scrutiny (including use of prior authorization by Medicaid in many U.S. states),[20] and there is a need for ongoing investigation into the comparative effectiveness and harms of available medications. Practice parameters for antipsychotic use produced by the American Academy of Child and Adolescent Psychiatry (AACAP) are referred to when assessing practice for pediatrics in the United States,[18] but these parameters may be considered outdated (all reviewed studies published prior to 2012) for providing the best evidence. This CER covers many psychiatric conditions, as well as behavioral issues, for which antipsychotics are being prescribed as mono- or adjunctive therapy, such that a diverse range of stakeholders can be provided with evidence on the relative benefits and harms of antipsychotics to make informed decisions.

Uses of Antipsychotics

The following sections describe the main features and uses of antipsychotics in the conditions covered by this CER.

Schizophrenia and schizophrenia-related psychosis. Schizophrenia and schizophrenia-related psychosis are grouped together because psychotic symptoms are prominent features of both conditions. The category includes schizophrenia, schizoaffective disorder, substance/medication-

induced psychotic disorder, or prodromal phase (ultra high-risk). Treatment of psychotic disorders or psychotic features includes long-term use of antipsychotic medications.

Autism spectrum disorders. Autism spectrum disorders include autism, pervasive developmental disorders, Asperger's disorder, and pervasive developmental disorders not otherwise specified.[3] These disorders are characterized by: 1) deficits in social communication and social interaction and 2) restricted repetitive patterns of behavior, interests, and activities. The U.S. National Health Interview Survey data indicated a four-fold increase in autism from 1997-1999 to 2006-2008.[21] This rising trend may be due to broadening diagnostic criteria, better ascertainment, and/or increased incidence.[22] Antipsychotics have been used to manage irritability or aggressive outbursts, reduce hyperactivity or repetitive behaviors, or promote sleep onset and continuity.[23]

Bipolar disorder. Bipolar disorder is a disorder characterized by unstable mood. There are several types of bipolar disorder: bipolar type I (manic episodes and depressive episodes occur independently), bipolar type II (hypomanic episodes and depressive episodes occur independently), cyclothymic disorder (episodes not meeting criteria for bipolar I or II), and (most prevalent) other or unspecific bipolar disorder (not meeting criteria for mania or hypomanic episodes in duration).[3] Antipsychotics may be used as the first-line medication, primarily for mania, even when psychosis is not present.

Attention-deficit/hyperactivity disorder and disruptive, impulse-control, and conduct disorders. Attention-deficit/hyperactivity disorder (ADHD) and disruptive, impulse-control, and conduct disorders are so named because the core symptoms disrupt the daily functioning of children and their families. The 2011/12 U.S. National Survey of Children's Health estimated that 11 percent of school-aged children have received a diagnosis of ADHD; this represents a 42 percent increase from 2003.[24] Antipsychotic medications are used predominantly to manage impulsive aggression and to help regulate negative emotions; they are also used in small doses to promote somnolence (an intended side effect), as many people with ADHD have sleep disturbance.

Obsessive-compulsive disorder. Obsessive-compulsive disorder (OCD) is a chronic condition characterized by obsessions (repetitive thoughts) or compulsions (repetitive behaviors) that cause distress and/or interfere with functioning. More than 90 percent of lifetime OCD diagnoses met the criteria for another psychiatric disorder.[25] Because of failure for many patients in response to first-line treatment with antidepressants and other therapies, treatment is often augmented with antipsychotics.[26]

Substance Use Disorder. The essential feature of a substance use disorder is a cluster of cognitive, behavioral, and physiological symptoms indicating that the individual continues using the substance despite significant substance-related problems.[3] Current craving is often used as a treatment outcome measure because it may be a signal of impending relapse. Substance use disorders are distinguishable from substance-induced disorders, such as substance-induced psychotic disorder, which will be captured in other categories in this CER. Dependence on some substances, particularly cocaine and psychostimulants, is related to their increase in release of the neurotransmitter dopamine.[27] Dopamine-related behaviors, including impulsivity, aggression, and sensation seeking, have also been shown to limit effectiveness of intensive outpatient therapies. Because of their blockade of dopamine transmission, antipsychotics may be used to reduce the reinforcing properties of these substances.[27] The use of antipsychotics in other cases, such as for alcohol use disorders, may in part rely on the dopamine-enhancing properties of some of these medications.[28]

Major and persistent depressive disorders, and disruptive mood dysregulation disorder. Of the depressive disorders, major depressive disorder (MDD) represents the classic condition. It is characterized by discrete episodes of at least 2 weeks duration, involving clear-cut changes in affect, cognition, and neurovegetative functions. Persistent depressive disorder requires symptoms of at least one year (two in adults). To address concerns about potential overdiagnosis and overtreatment of bipolar disorder in children, a new diagnosis, disruptive mood dysregulation disorder, is included for children up to age 18 years who exhibit persistent irritability and frequent episodes of extreme behavioral dyscontrol.[3] Antipsychotics are often used as adjunctive therapy for depressive disorders (i.e., aripiprazole, quetiapine, and olanzapine are indicated for treatment for major depression in adults), and have been shown to result in improvements in core symptoms of the condition.[29]

Anxiety disorder. Anxiety may occur in the course of another condition (e.g., bipolar, posttraumatic stress, OCD), but there are also several primary anxiety disorders (DSM-V does not classify OCD or posttraumatic stress disorder [PTSD] as anxiety disorders).[3] Prevalence rates of anxiety disorders (excluding rates for OCD and PTSD) in adolescence and in 18 to 29 year olds are substantial (21-25 percent from the National Comorbidity Surveys), and median age of onset in the adolescent sample was six. When onset is before adolescence, some disorders such as separation anxiety are more common; despite this, generalized anxiety disorder occurs in children and has a 12-month prevalence of 0.9 percent in the United States.[3] Antipsychotics may serve to augment treatment with antidepressants, and SGAs would be preferentially chosen over FGAs due to their effects at receptors contributing to antidepressive effects (e.g., serotonin, adrenergic).[30] Apart from anxiety symptoms, irritability and sleep disturbances are examples of symptoms which may be treated with antipsychotics.

Posttraumatic stress disorder. Posttraumatic stress disorder (PTSD) develops following a reaction of intense fear, helplessness, or horror resulting from a traumatic event.[31, 32] Characteristic symptoms include a persistent re-experience of the traumatic event (i.e., intrusions, flashbacks), persistent avoidance of stimuli associated with the trauma, numbing of general responsiveness, and persistent symptoms of increased arousal.[3] Individuals with PTSD often also experience psychotic symptoms such as paranoia, agitation, and delusional beliefs.[33] Median age of onset for a representative sample of adults in the United States' National Comorbidity Surveys was 23;[34] from National Survey of Adolescents data, the six-month prevalence of PTSD was estimated to be 3.7 percent in boys and 6.3 percent in girls.[35] Antipsychotics have been studied for use as monotherapy or adjunctive treatment (with antidepressants) for various symptoms in PTSD.[36, 37]

Eating disorders. Eating disorders are characterized by a persistent disturbance of eating or eating-related behavior that results in the altered consumption or absorption of food and that significantly impairs physical health or psychosocial functioning. Medications such as antidepressants, antipsychotics, and mood stabilizers may help treat anorexia nervosa and comorbid conditions (e.g., distorted thoughts) when given as part of a complete psychological treatment program.

Tic disorders. Tics are involuntary motor movements or vocalizations. Although some individuals have only motor or verbal tics, those with Tourette syndrome have both types. In most cases, Tourette syndrome is associated with co-morbid neuropsychiatric disorders—most commonly OCD or ADHD.[21, 38] Medications that inhibit dopamine reuptake, such as antipsychotics, generally help to reduce tics, but may induce tics in some cases.

Rationale for Update on this Topic

Comparative Effectiveness Reviews need to be regularly updated as new evidence is produced. Lack of attention to updating may lead to outdated and sometimes misleading conclusions that compromise health care and policy decisions. During updating, constituent elements of the originally formulated protocol (e.g., search strategy, eligibility criteria, and key questions) may be retained and sometimes extended/modified to accommodate newly identified evidence (e.g., new intervention, new outcome, or new subpopulation) or to address changing clinical or policy controversies.[39] A review of the original methodology may also find that different or newer approaches may improve the rigor in assessment of the results and the overall strength of evidence.

Although systematic reviews existed focusing on the efficacy of antipsychotics for specific indications, the previous CER published in 2012[1] included comparisons within and between classes of antipsychotics for a broad range of conditions. The strength of the evidence (SOE; indicating the degree of confidence in the findings across studies) was evaluated for a wide range of comparisons and outcomes; no outcomes were assessed as having high SOE, and few were evaluated as having moderate SOE (e.g., olanzapine caused more dyslipidemia and weight gain, but fewer prolactin-related events, than risperidone, and more weight gain than quetiapine). More studies providing evidence for the previous comparisons could change the SOE. The approval by the FDA of lurasidone, brexpiprazole, and cariprazine, and the return to market of molindone will add evidence. Moreover, the addition of several conditions of interest will expand the scope. Our preliminary update search has confirmed that there are at least 40 relevant trials either published or near completion.

There are also methodological considerations that may either change some conclusions of the original CER, or enhance our ability to inform decisions in some areas. The original assessment of SOE was frequently downgraded due to high risk of bias for the relevant studies, which included consideration of industry funding. Refinement in EPC program methods guidance on risk of bias assessments of individual studies, in particular in relation to the role of industry funding, may not lead to similar assessments in the updated review.[40] For some outcomes (especially harms which were evaluated across disorders), the use of mixed-comparison meta-analytical techniques (i.e., combining placebo and head-to-head trials across a variety of drug comparison) may be possible and allow for more quantitative assessment of differences between antipsychotics in the absence of many head-to-head trials. Moreover, the assessment of findings for patient and clinical subgroups relied upon within-study analyses which were highly variable and did not encompass harms data; applying analytical techniques with study-level data—although exploratory in nature[41]—would allow for examining the related key questions (KQ1a, b; KQ2 a, b) to a greater extent. Lastly, differences in some harms outcomes (e.g., weight gain and metabolic risks) have been shown to vary by condition,[42, 43] such that only using aggregate data on harms across conditions may not capture some information important for patient-level decision making.

Objective

This review will update the original review of 2012, with some modifications including expanding the indicated conditions, assessing harms by condition as well as across all conditions, and conducting additional analyses for capturing differential benefit and harm for subgroups based on patient and clinical characteristics. The findings will be useful for multiple

stakeholders, and inform efforts by professional societies to develop evidence-based recommendations and clinical practice guidelines to guide appropriate use in practice.

III.   Scope and Key Questions

Conditions of Interest

- Schizophrenia and schizophrenia-related psychoses, including schizoaffective disorder and substance/medication-induced psychotic disorder, or prodromic (ultra high-risk) psychosis.
- Autism spectrum disorders, including pervasive developmental disorder, autism, Rett's disorder, childhood disintegrative disorder, Asperger's disorder, and pervasive developmental disorder not otherwise specified.
- Bipolar disorder.
- Attention-deficit/hyperactivity disorder or disruptive, impulse-control, and conduct disorders, including conduct disorder, oppositional defiant disorder, intermittent explosive disorder, and other specified/unspecified disruptive, impulse-control, or conduct disorders.
- Obsessive-compulsive disorder.
- Substance use disorder.
- Major and persistent depressive disorders, or disruptive mood dysregulation disorder.
- Anxiety disorders.
- Posttraumatic stress disorder.
- Eating disorders (i.e., anorexia nervosa, bulimia nervosa, binge-eating disorder).
- Tic disorders (e.g., Tourette syndrome).
- Behavioral issues outside the context of a mental disorder, including aggression, agitation, anxiety, behavioral dyscontrol, irritability, mood lability, self-injurious behaviors, and insomnia.

Key Questions

For Each Condition of Interest

1. What are the benefits, in terms of intermediate and effectiveness outcomes, of first and second generation antipsychotics—at the level of individual antipsychotics and across each class—in comparisons with placebo, different doses of the same antipsychotic, or different antipsychotics in children and young adults (≤24 years)?
    (a) Do the benefits vary with respect to patient characteristics, such as age, sex, race/ethnicity, medical comorbidities, phase or features of disorder, and antipsychotic treatment history?
2. What are the benefits, in terms of intermediate and effectiveness outcomes, of first and second generation antipsychotics—at the level of individual antipsychotics and across each class—in comparisons with placebo, different doses of the same antipsychotic, or different antipsychotics in children and young adults (≤24 years)?
    (a) Do the benefits vary with respect to patient characteristics, such as age, sex, race/ethnicity, medical comorbidities, phase or features of disorder, and antipsychotic treatment history?
    (b) Do the benefits vary with respect to clinical characteristics such as dose of antipsychotic, or cotreatments including other antipsychotics, other medications, or nonpharmacologic therapy?
3. What are the harms of first and second generation antipsychotics—at the level of individual antipsychotics and across each class—in comparisons with placebo, different doses of the same antipsychotic, or different antipsychotics in children and young adults (≤24 years)?
    (a) Do the harms vary with respect to patient characteristics, such as age, sex, race/ethnicity, medical comorbidities, phase or features of disorder, and antipsychotic treatment history?
    (b) Do the harms vary with respect to clinical characteristics such as dose of antipsychotic, or cotreatments including other antipsychotics, other medications, or nonpharmacologic therapy?

Across All Conditions

4. Across all conditions of interest, what are the harms of first and second generation antipsychotics—at the level of individual antipsychotics and across each class—in comparisons with placebo, different doses of the same antipsychotic, or different antipsychotics in children and young adults (≤24 years)?
    (a) Do the harms vary with respect to patient characteristics, such as age, sex, race/ethnicity, medical comorbidities, phase of disorder, and prior exposure to antipsychotics?
    (b) Do the harms vary with respect to clinical characteristics such as dose of antipsychotic, or cotreatments including other antipsychotics, other medications or nonpharmacologic therapy?

IV. PICOTS-D Criteria

The PICOTS-D (patients, interventions, comparators, outcomes, timing, setting, and study design) applicable to this CER are presented below in Table 1. These criteria will guide all the stages of the CER, including literature searching and study selection, and data abstraction and analysis.

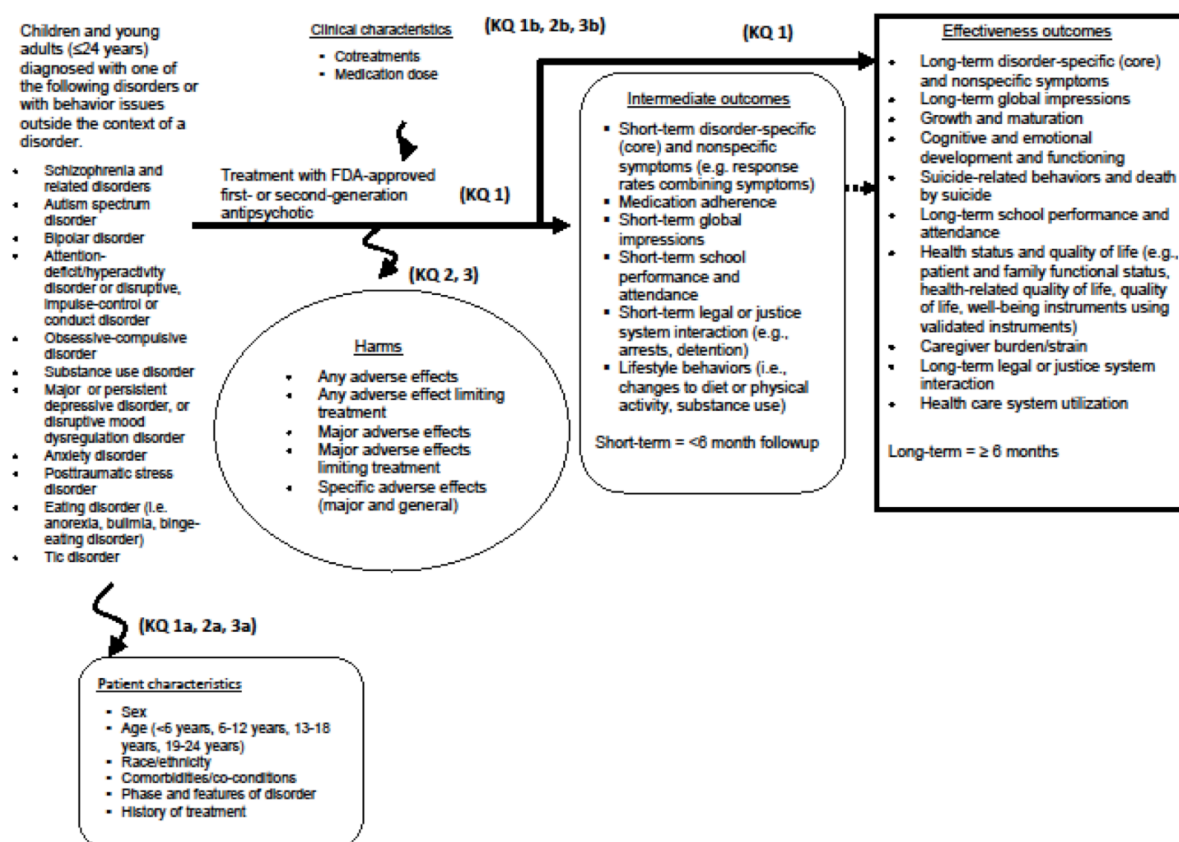**Table 1. PICOTS (population, interventions, comparators, outcomes, timing, setting)**

Category	Criteria
Population	Children and young adults (≤24 years) with one or more of the following conditions/issues: AD, ADHD, ASD, BD, DD, DICD, ED, OCD, PTSD, SUD, SZ, TD, or behavioral issues outside the context of a disorder (e.g., insomnia).
	Subpopulations based on patient characteristics: sex; age (<6 years, 6-12 years, 13-18 years, 19-24 years); race/ethnicity (i.e., percent nonwhite); comorbidities/co-conditions (e.g., ADHD, substance use); history of treatment (e.g., naïve, refractory); phase and features of disorder (e.g., acute mania vs. maintenance treatment [bipolar disorder], first-episode psychosis versus treatment in context of prior episodes [schizophrenia], presence of psychosis [disorders other than schizophrenia]).
Interventions	Any FDA-approved FGA (chlorpromazine, droperidol, fluphenazine, haloperidol, loxapine, molindone, perphenazine, pimozide, prochlorperazine, thiothixene, thioridazine, trifluoperazine)
	Any FDA-approved SGA (aripiprazole, asenapine, brexpiprazole, cariprazine, clozapine, iloperidone, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, ziprasidone)
	Subpopulations as per clinical characteristics: presence of cotreatments (e.g., other medication, nonpharmacological therapy, as reported); medication dose.
Comparators	Placebo/no treatment, any other antipsychotic, or same antipsychotic at different dose.
	Exclusion of non-antipsychotic medications as comparator.
Outcomes	KQ 1: intermediate and effectiveness outcomes (see lists in Methods section).
	KQ 2 & 3: any adverse effect (AE) and any major AEs; any or major AE limiting treatment (e.g., withdrawal due to AE); specific AEs (i.e., individual major or general AEs); persistence and reversibility of AEs (see list of major and general AEs in Methods)
Timing	No minimum followup duration
	Short term: <6 months
	Long term: ≥6 months-<12 months; 12 months+
Setting	Any setting
Study Design	Clinical trials (RCTs and NRCTs), controlled cohort studies (prospective or retrospective), controlled before-after studies (e.g., open-label extensions with comparator group, pooled analyses of individual patient-level data from one or a combination of similar trials).
Language	English

AD = anxiety disorder; ADHD = attention-deficit/hyperactivity disorder; ASD = autism spectrum disorder; BD = bipolar disorder; DD = major and persistent depressive disorder; DICD = disruptive, impulse-control, and conduct disorder; ED = eating disorder; FDA = Food and Drug Administration; FGA = first-generation antipsychotic; KQ = key question; NRCT = nonrandomized controlled trial; OCD = obsessive-compulsive disorder; PTSD = posttraumatic stress disorder; RCT = randomized controlled trial; SGA = second-generation antipsychotic; SUD = substance use disorder; SZ = schizophrenia or other related psychosis; TD = Tic disorders

V. Analytic Framework

Figure 1 is an analytic framework that depicts the structure used to address the Key Questions (KQs) for evaluating the benefits and harms of first- and second-generation antipsychotics in children and young adults (≤24 years of age). We will examine the benefits and harms of FDA-approved FGAs and SGAs in a population of children and young adults (≤ 24 years) diagnosed with one of the psychiatric conditions identified, or experiencing behavioral issues outside the context of a psychiatric diagnosis (e.g., sleep difficulties, agitation, aggression). In KQ1, benefit will be determined (by condition) for intermediate outcomes (e.g., short-term disorder-specific and nonspecific symptoms, short-term medication adherence, lifestyle behaviors), and effectiveness (e.g., long-term symptoms, growth and maturation, health status and quality of life, caregiver burden/strain) outcomes. In KQ2 we will assess harms within conditions in terms of medication-associated adverse effects categorized as major (e.g., mortality, development of diabetes) and general (e.g., extrapyramidal effects, weight gain, hyperprolactinemia). KQ3 will evaluate harms across all conditions. Within each KQ, we will assess outcomes for subgroups of patients or studies based on patient and clinical/treatment characteristics.

Figure 1. Analytic framework for the Key Questions (KQ) evaluating the comparative effectiveness of FDA-approved first and second generation antipsychotics in children and young adults ≤24 years old.

VI. Methods

The methods for this review of antipsychotics in children and young adults are based on the methods specified in the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (Methods Guide).[44]

A. Criteria for Inclusion/Exclusion of Studies in the Review

We will use the eligibility criteria outlined in the PICOTS-D as presented above under the Key Questions in Table 1. Additional details for the inclusion and exclusion criteria related to the PICOTS-D elements are presented here.

Population

Our population of interest is children (0-18 years) and young adults ≤24 years of age with psychiatric disorders or behavioral disturbances. Studies that enroll adults are included when at least 80 percent of patients are ≤ 24 years of age, or when subgroup analyses or individual data for patients within the eligible age range are provided. For KQs 1 and 2, we will exclude studies where the large majority (≥ 90 percent) of participants do not share a common diagnosis or behavioral issue as classified above in the Background section (e.g., study of patients with either schizophrenia or pervasive developmental disorder), unless subgroup analysis is reported for the separate conditions. One exception to this will be studies including patients experiencing psychosis (often first-episode) regardless of diagnosis; this group of studies is anticipated to be classified as a subgroup of studies within the schizophrenia and related psychosis condition category. For KQ 3 on harms across conditions, we will include studies where participants do not share a common diagnosis.

Interventions

The intervention drug must be an FDA-approved FGA or SGA (Table 2, Appendix A Tables A1 to A4) that is currently available in the United States; we will assess the evidence for all uses of these drugs, regardless of their approved indications. All formulations of drug delivery (e.g., tablet, liquid, injectable) and doses are eligible. Polypharmacy is common in these clinical populations; therefore, studies including patients taking other medications will be included providing these medications are monitored (dose) during the study. Likewise, studies of combination therapies (i.e., antipsychotic as an adjuvant to another nonantipsychotic drug and/or a nonpharmacological treatment) are acceptable providing that all study participants receive the same protocol in this regard (e.g., olanzapine and citalopram vs. olanzapine would be excluded).

Table 2. List of antipsychotics included in the CER

First generation antipsychotics	Second generation antipsychotics
• Chlorpromazine	• Aripiprazole
• Droperidol	• Asenapine
• Fluphenazine	• Brexpiprazole
• Haloperidol	• Cariprazine
• Loxapine	• Clozapine
• Molindone	• Iloperidone
• Perphenazine	• Lurasidone
• Pimozide	• Olanzapine

- Prochlorperazine
- Thiothixene
- Thioridazine
- Trifluoperazine
- Paliperidone
- Quetiapine
- Risperidone
- Ziprasidone

Comparators

Comparators of interest are placebo or no treatment, another antipsychotic, or the same antipsychotic at a different dose.

Outcomes

The intermediate and effectiveness outcomes listed below will be extracted as reported by study authors; for example, we will include relevant author-defined outcomes (such as percentage of participants gaining >7 percent body weight, or duration of response, remission, relapse, speed of response, reversibility of adverse effects, and non-compliance or withdrawal due to lack of efficacy/response, and time to discontinuation of medication) as long as these are accounting for benefit and harm outcomes of interest. We will also account for duration of response, that is, short- (<6 months) and long-term (≥6 months-<12 months; 12 months+) followup. Key outcomes that will be assessed for the strength of the body of evidence and considered for subgroup analyses are indicated by an asterisk; these key outcomes were chosen—using input from key informants and our Technical Expert Panel (TEP)—because they reflect outcomes most targeted by treatment with antipsychotics and of relatively high importance to patients, their families, and clinicians.

Intermediate outcomes

- Short-term disorder-specific (core) symptoms:
  ◦ Schizophrenia and related psychoses: positive* and negative symptoms*, disorganized behavior, impaired thought process, mood symptoms;
  ◦ Autism spectrum disorders: irritability (i.e., aggression, deliberate self-injury, and temper tantrums)*, qualitative impairment in social interactions*, communication*, restricted repetitive and stereotyped behaviors*, interests, and activities;
  ◦ Bipolar disorder: severity of mania*, anxiety, depression*, mood symptoms, psychotic features*;
  ◦ Attention-deficit/hyperactivity disorder or disruptive, impulse-control, and conduct disorders: aggression*, negativistic, hostile and defiant behavior, externalizing behaviors*, impulsivity*;
  ◦ Obsessive compulsive disorder: obsessive thoughts*, compulsive behavior*;
  ◦ Substance use disorder: cravings, abstinence/substance use days*;
  ◦ Major or persistent depressive disorder: depression*, irritability*, psychotic features (e.g., positive and negative symptoms)*;
  ◦ Anxiety disorder: anxiety*, irritability*;
  ◦ Posttraumatic stress disorder: hyperarousal*, avoidance behaviors*, intrusion*;
  ◦ Eating disorders: weight*, body mass index, cognitive distortions, eating disorder attitudes and beliefs;
  ◦ Tic disorders: motor and vocal tic frequency* and severity*;

- ◦ Behavioral issues outside the context of disorder or illness: aggression, agitation, anxiety, behavioral dyscontrol, irritability, mood lability, self-injurious behaviors, and sleep latency and duration.
- Short-term nonspecific or associated symptoms
  - ◦ Various (often composite or associated) psychiatric behaviors or symptoms (e.g., response rates*, anxiety in OCD, depression in tic disorders, sleep disorders, overall behaviors/symptoms in autism), and not including global assessments
- Medication adherence
- Short-term global impressions and functioning*
- Short-term school performance and attendance
- Short-term legal or justice system interaction (e.g., arrests, detention)
- Lifestyle behaviors (i.e., changes to diet or physical activity)

Effectiveness (patient- and family-important) outcomes

- Long-term (≥ 6 month followup) disorder-specific symptoms (see list above under Intermediate Outcomes)*
- Long-term (≥ 6 month followup) nonspecific or associated symptoms
  - ◦ Various (often composite or associated) psychiatric behaviors or symptoms (e.g., response rates*, anxiety in OCD, depression in tic disorders, sleep disorders, overall behaviors/symptoms in autism), and not including global assessments
- Long-term (≥ 6 month followup) global impressions and functioning*
- Growth and maturation
- Cognitive and emotional development and functioning*
- Suicide-related ideations or behaviors, or death by suicide*
- Long-term (≥ 6 month followup) school performance and attendance
- Occupational functional capacity
- Generic and specific health status and quality of life (i.e., patient and family functional status [e.g., social or relationship success, development of autonomy, and others tied to developmental level and family function], health-related quality of life, quality of life, well-being) using validated instruments*
- Caregiver burden/strain
- Long-term (≥ 6 month followup) legal or justice system interaction*
- Health care system utilization

Harms

Adverse effects (AEs) will be examined for each condition (KQ2) and across all conditions (KQ3). In addition to describing findings for each AE specified below, we will also analyze AEs in terms of: 1) any adverse event (AE) and any AE limiting treatment (i.e., non-compliance/withdrawal rates due to AEs), and 2) major AEs and major AEs limiting treatment.

Major adverse effects*

- Mortality
- Cerebrovascular disease-related events
- Development of diabetes mellitus
- Diabetic ketoacidosis

- Neuroleptic malignant syndrome
- Seizures
- Tardive dyskinesia
- Cardiomyopathies
- Cardiac arrhythmias
- Agranulocytosis

General adverse effects

- Neuromotor effects (e.g., extrapyramidal symptoms including dystonia, akinesia, akathisia)*
- Metabolic effects (e.g., metabolic syndrome, change in body composition [weight, BMI], fasting glucose, insulin sensitivity/resistance, dyslipidemia [total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides], blood pressure)*
- Prolactin-related effects and sexual dysfunction (e.g., hyperprolactinemia, AEs related to prolactin elevations [e.g., galactorrhea/bloody galactorrhea, hypogonadism], erectile dysfunction, infertility, oligo/amenorrhea, precocious puberty)*
- Agitation
- Constipation
- Somnolence* and fatigue
- Elevated transaminases
- Exercise intolerance
- Discontinuation syndrome (including symptoms related to motor [e.g., withdrawal-induced dyskinesias, dystonias], autonomic (e.g., disturbed temperature regulation, nausea] and psychoses [e.g., rebound psychosis]

Timing, Setting and Design

Studies published in 1987 or later will be included. There is no minimum duration of followup, although we will categorize most by short- (<6 months) or long-term (≥6 months-<12 months; 12 months+) followup. All settings are applicable.

Eligible study designs include controlled clinical trials (RCTs and NRCTs), controlled cohort studies (prospective or retrospective), and controlled before-after studies (e.g., open-label extensions of RCTs with a comparator group, pooled analyses of individual patient-level data from one or a combination of similar trials). We will not include case control studies as they are not an optimal study design for assessing causal inferences or measuring treatment effects. We will not include studies without comparators (e.g., noncontrolled cohort, before-after studies) for the same reason. Because this update includes additional conditions and study designs, we will need to re-screen citations from the original inclusion date of 1987.

Additional Criteria

We are including English-language publications because we believe it is unlikely that we will miss important data that are reported in non-English articles; effect sizes in language-restricted reviews have shown to not differ significantly (overestimating effect sizes by 2 percent) from those not having restrictions.[45] For all outcomes, we will include studies reported

by unpublished and published research articles, reports, and dissertations. Studies published exclusively in abstract form (e.g., conference abstracts) will not be included, but if relevant we will search for a complete report including contacting authors, as needed. For harms outcomes in KQ 2 and 3, we will also include studies unique to trial registries and regulatory agencies (i.e., not previously identified by other search strategies); these sources will also be used to identify unpublished data on harms from already included studies.

We do not have a minimum sample size for inclusion, nor do we have a minimum threshold for extent of incomplete followup or participant attrition; these factors will be considered during assessment of the strength of evidence (e.g., precision domain accounts for sample size across studies), and during sensitivity analyses in cases of substantial heterogeneity in findings in the data synthesis stage (see relevant sections).

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

The research librarian, in collaboration with the investigative team, will revise and implement the original search strategy to incorporate the changes to the conditions of interest. Because of the addition of several conditions, we will run all searches back to 1987 rather than 2010 as suggested for update searches.

We will conduct comprehensive searches in the following electronic databases: Ovid MEDLINE and Ovid MEDLINE® In-Process & Other Non-Indexed Citations, Cochrane Central Register of Controlled Trials via Cochrane Library, EMBASE® via Ovid, CINAHL Plus with Full Text via EBSCOhost, PsycINFO® via Ovid, ProQuest® Dissertations and Theses - Full Text, and TOXLINE via The National Library of Medicine. All searches will be restricted to English language studies published since 1987. Using a combination of controlled vocabulary and keywords, search filters for RCTs, NRCTs, and observational studies will be applied (where applicable) to the search results retrieved from the above listed databases.[46] Following submission of the draft report to AHRQ, searches will be re-run in the main databases (PubMed, EMBASE, PsycINFO, and CENTRAL) to identify any new publications. The search strategy will be developed in MEDLINE (Appendix B), peer reviewed by a second librarian, and adapted to accommodate the controlled vocabularies and search languages of the other databases.

Several other sources will be used for obtaining data from or reports of studies. Reference lists of relevant systematic reviews and guidelines (identified when searching bibliographic databases), and of included studies will be screened to identify potentially relevant (published or unpublished) studies. We will search ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform. We will search the conference proceedings of the following key scientific meetings for 2014-2015 to identify potentially relevant studies: American Academy of Child and Adolescent Psychiatry, International College of Neuropsychopharmacology, and International Society for Bipolar Disorders. We will also handsearch the Journal of Child and Adolescent Psychopharmacology and the Journal of the American Academy of Child and Adolescent Psychiatry (2014-2015). Drug manufacturers and other relevant stakeholders (via AHRQ's Scientific Resource Center) will be notified of the opportunity to submit scientific information relevant to the interventions of this systematic review. We will contact authors (by email with three attempts) of relevant protocols, trial registries, and abstracts that identify studies not located in the searches to obtain any reports or publications of completed studies. We will search for clinical study reports from the European Medicines Agency, and search Drugs@FDA for Medical/Clinical and Statistical review

documents; as for the original CER, we will only search regulatory documents containing data for patients 18 years of age or younger. This decision is because of the (1) complexities of determining the age of study participants referenced in regulatory reports, and (2) feasibility issues arising when considering the number of drugs being evaluated. Where reviews that are anticipated to have data from trials (1987 or later) in pediatrics are not available we will request these. Any studies suggested by Peer Reviewers or through public comment on the draft report will also be assessed for eligibility; eligible studies will be incorporated into the final report.

All results of the database searches will be imported into an EndNote® database (Thomson Reuters, New York, NY). Results from other searches will be documented in a Microsoft Excel database (Microsoft Corp., Redmond, WA). We will track the screening and selection results in EndNote.

C. Study Selection

For the database searches, two reviewers will independently screen the titles and abstracts (when available) using broad inclusion/exclusion criteria. Citations will be classified as "include/unsure," "exclude," or "reference" (i.e., conference abstracts, protocols, and systematic reviews). One reviewer will review the "reference" group and will conduct all other searching as outlined in the above section. The full text of all studies classified as "include/unsure" or identified after reviewing the reference citations will be retrieved for full review; two reviewers will independently assess eligibility using a standard form that outlines the inclusion and exclusion criteria. Disagreements on final inclusion of all studies will be resolved through consensus or third party adjudication. We will revise the inclusion/exclusion form from the original review (e.g., add conditions and new drugs) and pre-test the form on a sample of studies. The title/abstract screening and full-text selection processes will be conducted and documented in Endnote, with exportation to Excel for the purposes of comparison for consensus procedures.

D. Data Abstraction and Data Management

Data extracted from the original CER was retrospectively added into the Systematic Review Data Repository™ (SRDR; http://srdr.ahrq.gov/).[47] We will use the Excel form created for this purpose for data extraction of the new studies, but will revise the form as required to accommodate the changes in this update. The data extraction form includes elements relevant to the Key Questions, including population characteristics, study characteristics (including funding source), descriptions of the intervention(s) and comparator(s)—including dose, route of administration, etcetera—analytic details including subgroup analysis on treatment modification, and outcomes including outcome type, timing and definitions. Any additional data required from the original studies (e.g., to capture new subgroup data) will be also extracted. The modified form will be pilot tested using a sample of the original CER's studies by all review team members involved in data extraction and analysis. One review team member will extract data for each study, and a senior level team member will verify all data extraction.

As done for the original CER, when there are multiple publications associated with a study we will consider the earliest report of the main (primary) outcome data to be the primary data source. We will extract data from the primary source first and then add outcome data reported in the secondary/associated publications and data sources (e.g., FDA reports). We will reference the primary source throughout the evidence report; all associated literature will be tabulated for reference. When the study design changes during the process of a research program, for example

from an RCT phase to a controlled, observational phase, we will consider these two different studies.

We will classify studies that directly compare one antipsychotic with another antipsychotic, or two different doses/administration of the same antipsychotic, as "head-to-head" studies and studies that compared an antipsychotic with placebo or no treatment as "placebo" studies. Studies with three or more treatment groups could provide data both for "head-to-head" and "placebo" comparisons.

We will report outcomes only if quantitative data are reported or can be derived/estimated from graphs (using the measurement tool of Adobe Acrobat 9 Pro [Adobe Systems Inc., California, U.S.]).We will not include outcomes that are only described qualitatively (e.g., "there was no difference between the groups") or reported only as a p-value in the data analysis. We will follow the decisions made during the original review (via consultation with clinical experts) regarding which outcome measurement tools to use for short- and long-term disorder-specific and nonspecific symptoms in KQ 1; the list (Appendix C) will be modified to capture outcomes specific for the newly added conditions (e.g., substance use) or behavioral issues, as needed. We will extract the total/composite score for each instrument measurement, when these align well with our outcomes of interest; for some tools, only the subscores/domains will be used because they may best align with our key disorder-specific outcomes.

We will record intention-to-treat results, if possible. For continuous outcomes measures, we will extract (by arm) the mean baseline and endpoint or change scores, standard deviations (SD) or other measure of variability, and number analyzed. We will not include outcome data from studies that did not provide a followup change or endpoint mean or data that could be used to calculated followup scores. If necessary, we will approximate means by medians. If standard deviations are not given, they will be computed from p-values, 95% confidence intervals (95% CIs), z-statistics, or t-statistics. If computation is not possible they will be estimated from upper bound p-values, ranges, inter-quartile ranges, or (as a last resort) by imputation using the largest reported SD from the other studies in the same meta-analysis. When computing SDs for change from baseline values, we will assume a correlation of 0.5, unless other information is present in the study that allows us to compute it more precisely. For dichotomous outcomes, we will report counts or proportions, and sample size, by study arm. We will extract data in terms of short- (<6 month) and long-term (≥6 month - <12 months; 12 months+) followup; when there are more than one timepoints in a study within each of these strata we will use the longest followup duration.

For AEs we will follow the original CER decisions, based on monitoring guidelines proposed by Correll et al.,[17] for determining which data to extract for each of the categories specified in Key Questions 2 and 3 (see Table 3). Adverse effects will be recorded for each study group including the description of the AE, how it was identified, defined and measured, the number of patients in each group, and the number of patients affected. For continuous adverse event measures (e.g., weight or prolactin levels), we will extract the mean change or endpoint score, SD, and sample size. For dichotomous data, each event will be counted as if it represents a unique individual; because a single individual might experience more than one AE, this assumption may overestimate the number of people having an AE. Only numerical data for AEs will be extracted; that is, we will make no assumptions on lack or presence of an AE if this is not reported, and studies that report only p-values or report one arm to have fewer events than another will not be included. By only using reported data, we may underestimate the number of patients for whom a particular adverse event is observed. We will extract data (taking care to

avoid duplication with other study reports) on harms from trial registries and regulatory agency reports of pediatric trials.

For our outcomes of "any" AE or "major" AE limiting treatment, we will use the data reported by the authors, as applicable and after consultation with clinical experts. If the incidences of AEs limiting treatment are presented in terms of specific AEs (e.g., withdrawals due to X, Y, Z, etc.), we will group these into "any" AE (i.e., general and major) and "major" AE according to our classification. For each AE, we will report the number of studies that provide data for the AE. We will also report summary totals of the number of individuals in the medication groups who are reported to have experienced the event and the total number of patients in the medication groups in relevant trials. The dose of each medication that is associated with an adverse event will be recorded to facilitate interpretation.

Data on within-study subgroup analysis will be collected, including: subgroups (independent variables), the type of analysis (e.g., subgroup/stratified or regression analysis), the outcomes assessed (dependent variables), and the authors' conclusions. We will collect data suitable for all patient and clinical characteristics for performing our own subgroup analyses based on study-level data (see Section F).

All data used in the quantitative analyses will be deposited into SRDR at the completion of the review.

Table 3. Adverse event outcome data for extraction in the comparative effectiveness review

Adverse event categories	Specific outcome data extracted
Mortality	ND
Cerebrovascular events	ND
Weight and body composition	Weight, percent with ≥7% weight change, weight status (e.g., % normal, overweight, etc.), BMI, BMI percentiles, fat mass, waist circumference
Dyslipidemia	Incidence of dyslipidemia, total cholesterol, LDL and HDL cholesterol, triglycerides, ratio of triglycerides to HDL cholesterol
Insulin resistance and diabetes	New-onset diabetes, exacerbation of previous diabetes, diabetic ketoacidosis, metabolic syndrome, HbA1c, glucose, insulin, HOMA–IR
Prolactin-related and sexual	Amenorrhea, oligomenorrhea, erectile dysfunction, decrease libido, hirsutism, breast symptom, galactorrhea, prolactin levels
Neuromotor	EPS scales, akathisia, tardive and withdrawal dyskinesia, dystonia
Cardiac	MI, cardiomyopathies, myocarditis, arrhythmias, abnormal ECG, QTc interval, hypertension, hypotension, orthostasis, postural hypotension, blood pressure, pulse, heart rate
Sedation	Sedation, somnolence, fatigue, tiredness
Liver toxicity	Liver damage, liver function test, liver enzyme levels (AST, ALT, GGT)
Neutropenia and agranulocytosis	Incidence of neutropenia, incidence of agranulocytosis, WBC counts
Thyroid dysfunction	Serum total thyroxine, serum free thyroxine, TSH
Seizures	ND
Neuroleptic malignant syndrome	ND
Constipation	ND
Exercise intolerance	ND
Precocious puberty	ND

ALT = alanine transaminase; AST = aspartate transaminase; BMI = body mass index; ECG = electrocardiogram; EPS = extrapyramidal symptoms; GGT = gamma-glutamyl transpeptidase; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; HOMA–IR = homeostasis model assessment of insulin resistance; LDL = low-density lipoprotein; MI = myocardial infarction; ND = not described (denotes categories that were not further subcategorized); QTc = QT interval corrected for heart rate; TSH = thyroid stimulating hormone; WBC = white blood cell

E. Assessment of Methodological Quality of Individual Studies

Two experienced reviewers will independently assess the methodological quality of all original and new studies and resolve discrepancies through consensus. We will re-assess original studies because of changes to guidance in the EPC program made subsequent to the original CER, and because of the addition of harms-specific assessment tools. Different tools will be employed based on study design and outcomes (e.g., subjective versus objectively measured effectiveness outcomes, and for harms).

For RCTs and NRCTs we will use the Cochrane Collaboration Risk of Bias tool,[48] with some modification based on EPC Methods guidance.[44] This tool consists of six domains (sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and "other" sources of bias) and a categorization of the overall risk of bias. We will not consider funding source as a source of "other" bias; the main contributors to bias from industry-funded studies are related to publication bias, and selective outcome and analysis reporting, which will be assessed separately by outcome during the assessment of the strength of evidence (publication/reporting bias domain)—see below section G. We will also not assess selective outcome reporting at the study level, since this will be considered within the reporting bias domain of our assessment of the strength of evidence for individual outcomes across studies.[49] Blinding and incomplete outcome data will be assessed separately for subjective outcomes (e.g., quality of life or function scales) and objective clinical outcomes (e.g., weight gain, glucose tolerance test). The overall assessment is based on the responses to individual domains. If one or more individual domains are assessed as having a high risk of bias, the overall score will be rated as high risk of bias. The overall risk of bias will be considered low only if all components are rated as having a low risk of bias. The risk of bias for all other studies will be rated as medium. Information will be collected for each study on the source of funding.

For cohort studies, we will use the Newcastle-Ottawa Quality Assessment Scale.[50] The scale comprises seven items that evaluate three domains of quality: sample selection, comparability of cohorts, and assessment of outcomes. Each item that is adequately addressed is awarded one star, except for the "comparability of cohorts" item, for which a maximum of two stars can be given. The overall score is calculated by tallying the stars. We considered a total score of 6 to 8 stars to indicate high quality, 4 or 5 stars to indicate moderate quality, and 3 or fewer stars to indicate poor quality.

F. Data Synthesis

For each condition we will summarize the characteristics of included studies qualitatively and present important features of the study populations, study designs, interventions, comparators, and reported outcomes in summary tables. For each KQ, we will synthesize data in the following order based on type of comparison (as possible depending on data): individual FGAs vs. SGAs, aggregate (across class) data for FGAs vs. SGAs, within-class comparisons between individual FGAs and individual SGAs (other drug or dose), and then individual and aggregate data for FGAs vs. placebo and SGAs vs. placebo.

For each outcome we will define a minimum important difference (MID; i.e., the smallest difference between groups that could lead to a change in management). For quality of life and other continuous patient/parent/clinician-reported measures, we will use an MID of one-fifth standard deviation (0.20 SD using data from pooled studies, or 0.2 standardized mean difference [SMD] if pooling a group of conceptually similar outcomes) which corresponds to a small effect size. For binary outcomes, we will use a relative risk (RR) of 0.75 and 1.25 to represent

clinically important benefit and harm, as suitable to the measure.[51] For other outcomes, we will seek input, if needed, from the TEP; to minimize potential bias when obtaining input for determining MID thresholds we will provide information on the reported outcome definitions (e.g., units, scales etc.) but not the findings from studies. Author-defined outcomes for "treatment response" typically use values similar to our MID (e.g., 20-30 percent improvement on included scales), such that we will interpret statistically significant differences in response rates to be clinically significant.

In general, we will combine results from studies when there is sufficient clinical (i.e., population characteristics, interventions, outcome ascertainments) and methodological (i.e., study design, conduct and quality) similarity. For example, within each category of conditions, we will only combine data on outcomes from different diagnoses (e.g., different anxiety disorder diagnoses) if there is sufficient homogeneity in the patient and clinical characteristics, including drug dosage. We will often start with combing all studies within a condition category and then use our a priori defined list of patient and intervention subgroups to explore the heterogeneity. We will combine results from RCTs with NRCTS, but not with cohort studies for which we will not pool results. In addition to the review team's assessment of clinical and methodological heterogeneity, we will also consider heterogeneity and consistency depending on the approach used for each analysis, as appropriate.

Various approaches to synthesizing the evidence are available including direct pairwise meta-analysis and methods that combine direct and indirect evidence (i.e., network meta-analysis or mixed treatment comparisons).[52-54] Moreover, approaches can vary by whether or not multiple outcomes, and/or covariates (e.g., age of patients, drug dose) are incorporated. The summary effect from direct comparisons (e.g., an SGA versus placebo, one SGA or FGA versus another SGA or FGA) for one outcome (using mean difference or SMD) at a similar timepoint is meaningful as a first approach. The original review focused on these pairwise meta-analyses. However, such an approach does not allow for comparisons between drugs that may not have much direct evidence (e.g., drug A has been compared to drug B and C, but drugs B and C have not been directly compared), and is limited by the number of studies reporting on the appropriate outcome. We anticipate that a multivariate, network meta-analyses may be possible for one or more outcomes within some conditions (e.g., symptoms in schizophrenia), and/or when comparing harms across all conditions. Meta-regression analysis—using dichotomous or continuous variables representing patient or clinical characteristics—may also be a viable option for exploring effectiveness or harms for subgroups based on study-level data. In the event that studies cannot be pooled, evidence tables will be produced and a narrative summary of the results will be presented.

Pairwise Meta-analysis

For pairwise meta-analyses, we will employ a Bayesian random effects model using WinBUGs software.[55, 56] We will only combine data from different outcome scales (e.g., aggression scales or subscales) if the scales are considered to measure the same construct; we will consult our Technical Experts on an as needed basis. We will report pooled MD, SMD, or RR with corresponding 95 percent credibility intervals (95% CrI). We will use deviance statistics to comment of the heterogeneity of the results. The decision to pool studies will not be based on the deviance statistics, but rather on interpretation of the clinical and methodological differences between studies. When substantial heterogeneity is suspected, we will conduct sensitivity analyses if appropriate (e.g., in the presence of studies with outlying effect sizes, for studies rated as high risk of bias in some domains such as incomplete [<70 percent] outcome data or lack of

allocation concealment, parallel versus cross-over designs). Heterogeneity will also be examined during our planned subgroup analyses. Where there are at least eight studies in a meta-analysis, we will analyze publication bias both visually using the funnel plot and quantitatively using Egger's test.[57]

Network Meta-analysis

Since we are interested in comparisons within and across classes of FGAs and SGAs, approaches to consider inferences from indirect data are suitable. Rather than providing a simple pairwise analysis of similar comparisons (e.g., a group of interventions versus usual care) through standard meta-analysis, a network meta-analysis allows for simultaneous evaluation of a suite of comparisons while still preserving the within-study randomization. A network of different comparisons is constructed (with "nodes" representing the different medications) to consider both direct evidence from comparisons of similar interventions/nodes and indirect evidence from comparisons where one intervention is in common, but not all (e.g., intervention A vs. placebo, and intervention B vs. placebo infer knowledge about intervention A vs. intervention B).

In addition to multiple comparisons, meta-analytical approaches have been developed to incorporate multiple outcomes and these can be incorporated into the network meta-analysis.[58, 59] As in the original review, the studies will report several outcomes with variability in overlap between studies, and between outcome metrics. One example is when most, but not all, studies report a composite score for a specific measurement tool but others only report one subscale or report on a different tool measuring a similar group of symptoms. A multivariate approach allows for the borrowing of strength across the entire set of relevant studies, and also enables the correlation between outcomes (both within and between studies) to be directly estimated. This allows individual studies to contribute partial information, and for missing outcomes to be readily imputed or predicted. For conditions where this approach might be possible (e.g., schizophrenia had the most studies in the original report), we will plan to tabulate all key effectiveness outcomes reported by the relevant studies and choose the outcomes reported by the most studies, to a maximum of four outcomes. For harms we will concentrate on outcomes related to three clusters of harms (each potentially serving a separate analysis): 1) metabolic effects (e.g., weight, BMI, dyslipidemia, fasting blood glucose), 2) neuromotor effects (e.g., dystonia, extrapyramidal syndrome, akathisia), and 3) hyperprolactinemia and sexual dysfunction (e.g., erectile dysfunction, precocious puberty); our ability to investigate each of these clusters will depend on the extent of reporting in enough studies to create a network of direct and indirect comparisons. Any decisions to combine outcomes within these clusters (e.g., weight and BMI changes as an SMD) will be made with input from clinical/content experts, who will be presented with data on outcome descriptions but blinded to study characteristics (e.g., authors) and results.

A key assumption underlying the validity of combining direct and indirect evidence is that both evidence types are consistent with one another. For example, should indirect evidence favor treatment A over treatment B (via respective direct comparisons with treatment C), then the corresponding direct evidence should support the same conclusion. We can reduce the likelihood of inconsistency by ensuring that the studies included in the meta-analysis are as similar as possible with respect to the underlying populations and implementation of treatments, and by explicitly modeling factors that might induce inconsistency. For the latter, we may be able to model factors thought to modify treatment effectiveness, or harm, as outlined in subgroups of

interest using study-level data on age, previous exposure to antipsychotics, drug dose, and so forth.[60]

When using this Bayesian multivariate, network meta-analysis approach, all unknown parameters will be given weakly-informative prior distributions and will be estimated using Markov chain Monte Carlo methods in WinBUGS software. The model will be run for 100,000 iterations, with the first 90,000 samples conservatively discarded as burn-in, leaving 10,000 for inference. We will conduct convergence diagnostics and assess the fit of the model, and the analysis will be checked for consistency by contrasting direct and indirect estimates. We will report the results of our convergence diagnostics and discuss the implications as applicable (e.g., less confidence in estimated effects in cases of inconsistency). We will obtain estimates of the treatment effects and rank probabilities for each treatment strategy (e.g., probability that a particular drug is the "best" for a particular outcome).

Analysis of Subgroups

Our primary approach to answer parts (a) and (b) of each KQ will be to record any within-study subgroup analyses performed by study investigators using individual patient data; these results preserve the within-study randomization. Because these results are often based on diverse methodology and may be difficult to interpret across the body of evidence, we will also perform our own subgroup analyses using study-level data, as possible, using formal statistical approaches (e.g., using meta-regressions, or incorporating data from one or more patient or clinical characteristics as variables in the Bayesian models used for the pairwise meta-analysis) or by stratifying the results of the pairwise meta-analyses by subgroup variables. When determining whether entire studies fall into a particular subgroup category (e.g., first-episode psychosis), we will consider ≥80 percent of the study population meeting the criteria as sufficient. The approach to choosing outcomes for inclusion is anticipated to be similar to that for the network meta-analysis as described above, but may rely on one common outcome depending on the analytical approach undertaken—for benefit outcomes we will choose the key outcome reported by the most studies; for harms we will focus on the three clusters of harms outlined in the preceding section. Ideally, we would choose outcomes within the effectiveness outcomes and major harms categories, although it is anticipated based on the results of the original CER that there will be insufficient data for these outcomes. We will employ regression analyses when: for continuous variables (e.g., age, duration of treatment) there are at least six to ten studies reporting on the outcome within a specific subgroup, and for categorical variables (e.g., first-onset psychosis versus prior episodes) there are at least three studies for each category level. The number of sufficient studies serves as a rule of thumb for the lower bound that investigators can consider for a meta-regression, but power will vary according to the size and variability of the effect. Similar to the network approach, these analyses would rely on study-level data, such that the results would be considered observational in nature.

G. Grading the Strength of the Body of Evidence

We will follow the Methods Guide and updated guidance[49] to evaluate the strength of the body of evidence (SOE) for the key outcomes and comparisons identified in the section on Inclusion/Exclusion criteria (p.13). For these assessments, we will largely rely on evidence from direct rather than mixed comparisons (e.g., data from network meta-analysis); methods for assessing the SOE for mixed comparisons (particularly with respect to the weight of the direct vs. indirect data) are still at a theoretical stage. The body of evidence will be graded by one

reviewer, and reviewed by a second reviewer. Disagreements will be resolved through discussion or by consulting with a third reviewer, as needed. Tables of findings will be generated for all outcomes and comparisons that have greater than insufficient SOE; the outcomes having insufficient SOE will be described in the text. We will also provide a summary of the differences in SOE assessments between the original CER and this update.

Trials and observational evidence will be graded separately for each outcome-comparison pair, with the overall SOE incorporating both study designs. As a starting point the SOE is assigned as high for evidence from trials, and low (benefit outcomes) or moderate (harm outcomes, being less influenced by bias) for evidence from observational studies. Thereafter we will examine and potentially downgrade the SOE based on five core domains: study limitations, consistency, directness, precision, and reporting bias. For outcomes where there is evidence from high-quality observational studies, we will also consider the additional domains of dose-response association, plausible confounding, and strength of association (i.e., magnitude of effect), to potentially upgrade the SOE; this will only be undertaken if serious limitations have not already been shown when assessing the trial or observational evidence for the other domains.[61]

We define the *study limitations* domain (low, medium, or high) on the basis of methodological quality. RCTs and NRCTs may be downgraded one or two levels depending on the numbers of trials assessed as having high risk of bias. Evidence from observational studies will be downgraded when studies have moderate or poor quality. We will rate *consistency* (consistent, inconsistent, unknown [if there is only one study]) by assessing the magnitude of the effects of the included studies (e.g., inconsistent when lack of overlap in 95% CrIs for most studies). We will assess *directness* of the evidence (direct or indirect) on the basis of the use of surrogate outcomes (when assessing our benefit outcomes), the need for indirect comparisons, or the evaluation of drug doses highly uncommon in clinical practice. We will assess *precision* (precise or imprecise) on the basis of sample size and, if size is adequate, the degree of certainty surrounding the effect estimate. For outcomes where thresholds of MID are used/determined, we will downgrade this domain once if the pooled 95% CrI crosses both no difference (0 MD or 1.0 RR) and our MID threshold, and twice when the 95% CrI crosses thresholds both for and against the intervention. For continuous outcomes, more than 400 total enrolled patients will generally be considered to offer precise data based on adequate power to detect a 0.2 standardized effect size. For binary outcomes, the sufficiency of the sample size will be mainly based on event rates in the control group.[62] A precise estimate is one that allows for a clinically useful conclusion. *Reporting bias* (suspected or undetected) will be evaluated with respect to publication bias, selective outcome reporting bias, and selective analysis reporting bias; this will only be assessed for evidence from trials, and when the overall SOE has not already been downgraded to insufficient based on other domains. Where there are at least ten studies having variable sample sizes in a meta-analysis, we will analyze publication bias both visually using the funnel plot and quantitatively using Egger's test.[58] For selective reporting and analysis biases, we will evaluate results across studies qualitatively on the basis of completeness of reporting for individual studies and reporting patterns across studies. To assist with publication/reporting bias assessments, we will seek abstracts (publication bias only), study protocols, registries, and studies/data from regulatory documents (for publication and outcome reporting bias).[63]

We will rate the body of evidence for each outcome and comparison using four SOE grades which indicate our level of confidence that the evidence reflects the true (direction or/and magnitude of) effect for the major comparisons of interest:

High   **We are very confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has few or no deficiencies. We believe that the findings are stable, i.e., another study would not change the conclusions.

Moderate   **We are moderately confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.

Low   **We have limited confidence that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.

Insufficient   **We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome**. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

H. Applicability

We will assess the applicability of the findings with respect to our PICOTS elements. We will summarize common features of the study populations and documented diagnoses. We will consider patient ages, intervention settings, treatment histories, co-occurring diagnoses, and symptom severity reported in the included studies and the degree to which the populations studied reflect the target populations for practice.

VII.   References

1.   Seida JC, Schouten JR, Mousavi SS, et al. First- and Second-Generation Antipsychotics for Children and Young Adults. Comparative Effectiveness Review No. 39. (Prepared by the University of Alberta Evidence-based Practice Center under Contract No. 290-2007-10021.) AHRQ Publication No. 11(12)-EHC077-EF. Rockville, MD. Agency for Healthcare Research and Quality. February 2012.

2.   Kondo K, O'Neil, M., Mathis, S., et al. Surveilllance Report: First- and Second-Generation Antipsychotics for Children and Young Adults. Agency of Healthcare Research and Quality; 2014. http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1571.

3.   American Psychological Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth edition. Washington, DC: American Psychiatric Association, 2013.

4.   Olfson M, Crystal S, Huang C, et al. Trends in antipsychotic drug use by very young, privately insured children. J Am Acad Child Adoles Psychiatr. 2010 Jan;49(1):13-23. PMID: 20215922.

5.   Pathak P, West D, Martin BC, et al. Evidence-based use of second-generation antipsychotics in a state medicaid pediatric population, 2001-2005. Psychiatric Services (Washington, DC). 2010 Feb;61(2):123-9. PMID: 20123816.

6.   Zito JM, Safer DJ, De Jong-Van Den Berg LT, et al. A three-country comparison of psychotropic medication prevalence in youth. Child Adolesc Psychiatr Ment Health. 2008;2(1):26. PMID: 18817536.

7.   Zito JM, Safer DJ, Dosreis S, et al. Psychotropic practice patterns for youth: a 10-year perspective. Arch Pediatr Adolesc Med. 2003 Jan;157(1):17-25. PMID: 12517190.

8.   Zito JM, Safer DJ, Sai D, et al. Psychotropic medication patterns among youth in foster care. Pediatrics. 2008 Jan;121(1):e157-63. PMID: 18166534.

9.   Olfson M, King M, Schoenbaum M. Treatment of young people with antipsychotic medications in the United States. JAMA Psychiatry. 2015;72(9):867-74. PMID: 26132724.

10. Zito JM, Derivan AT, Kratochvil CJ, et al. Off-label psychopharmacologic prescribing for children: History supports close clinical monitoring. Child Adolesc Psychiatr Ment Health. 2008;2(1):24. PMID: 18793403.

11. IMS Institute for Healthcare Informatics. The Use of Medicines in the United States: Review of 2010. NJ, USA: IMS Institute for Healthcare Informatics, 2011.

12. Crystal S, Olfson M, Huang C, et al. Broadened use of atypical antipsychotics: safety, effectiveness, and policy challenges. Health Affairs (Project Hope). 2009 Sep-Oct;28(5):w770-81. PMID: 19622537.

13. Patten SB, Waheed W, Bresee L. A review of pharmacoepidemiologic studies of antipsychotic use in children and adolescents. Can J Psychiatry. 2012 Dec;57(12):717-21. PMID: 23228229.

14. Fraguas D, Correll CU, Merchan-Naranjo J, et al. Efficacy and safety of second-generation antipsychotics in children and adolescents with psychotic and bipolar spectrum disorders: comprehensive review of prospective head-to-head and placebo-controlled comparisons. Euro Neuropsychopharmacol. 2011 Aug;21(8):621-45. PMID: 20702068.

15. Panagiotopoulos C, Ronsley R, Elbe D, et al. First do no harm: promoting an evidence-based approach to atypical antipsychotic use in children and adolescents. J Can Acad Child Adolesc Psychiatr. 2010 May;19(2):124-37. PMID: 20467549.

16. Pringsheim T, Lam D, Ching H, et al. Metabolic and neurological complications of second-generation antipsychotic use in children: a systematic review and meta-analysis of randomized controlled trials. Drug Safety. 2011 Aug 1;34(8):651-68. PMID: 21751826.

17. Correll CU, Penzner JB, Parikh UH, et al. Recognizing and monitoring adverse events of second-generation antipsychotics in children and adolescents. Child Adolesc Psychiatr Clinics NA. 2006 Jan;15(1):177-206. PMID: 16321730.

18. Rettew DC, Greenblatt J, Kamon J, et al. Antipsychotic medication prescribing in children enrolled in Medicaid. Pediatrics. 2015 Apr;135(4):658-65. PMID: 25733747.

19. Birnbaum ML, Saito E, Gerhard T, et al. Pharmacoepidemiology of antipsychotic use in youth with ADHD: trends and clinical implications. Curr Psychiatr Rep. 2013 Aug;15(8):382. PMID: 23881713.

20. Schmid I, Burcu M, Zito JM. Medicaid prior authorization policies for pediatric use of antipsychotic medications. JAMA. 2015 Mar 3;313(9):966-8. PMID: 25734740.

21. Perou R, Bitsko RH, Blumberg SJ, et al. Mental health surveillance among children--United States, 2005-2011. MMWR Surv Sum. 2013 May 17;62 Suppl 2:1-35. PMID: 23677130.

22. Baird G, Simonoff E, Pickles A, et al. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the special needs and autism project (SNAP). Lancet. 2006 Jul 15;368(9531):210-5. PMID: 16844490.

23. Coury DL, Anagnostou E, Manning-Courtney P, et al. Use of psychotropic medication in children and adolescents with autism spectrum disorders. Pediatrics. 2012 Nov;130 Suppl 2:S69-76. PMID: 23118256.

24. Visser SN, Danielson ML, Bitsko RH, et al. Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United states, 2003-2011. J Am Acad Child Adolesc Psychiatr. 2014 Jan;53(1):34-46 e2. PMID: 24342384.

25. Ruscio AM, Stein DJ, Chiu WT, et al. The epidemiology of obsessive-compulsive disorder in the national comorbidity survey replication. Mol Psychiatr. 2010 Jan;15(1):53-63. PMID: 18725912.

26.   Diniz JB, Shavitt RG, Fossaluza V, et al. A double-blind, randomized, controlled trial of fluoxetine plus quetiapine or clomipramine versus fluoxetine plus placebo for obsessive-compulsive disorder. J Clin Psychopharm. 2011 Dec;31(6):763-8. PMID: 22020357.

27.   Kishi T, Matsuda Y, Iwata N, et al. Antipsychotics for cocaine or psychostimulant dependence: systematic review and meta-analysis of randomized, placebo-controlled trials. J Clin Psychiatr. 2013 Dec;74(12):e1169-80. PMID: 24434105.

28.   Vergne DE, Anton RF. Aripiprazole: a drug with a novel mechanism of action and possible efficacy for alcohol dependence. CNS Neuro Dis Drug Tar. 2010 Mar;9(1):50-4. PMID: 20201815.

29.   Nelson JC, Mankoski R, Baker RA, et al. Effects of aripiprazole adjunctive to standard antidepressant treatment on the core symptoms of depression: a post-hoc, pooled analysis of two large, placebo-controlled studies. J Affect Disord. 2010 Jan;120(1-3):133-40. PMID: 19656577.

30.   Chen YC, Chen CK, Wang LJ. Quetiapine fumarate augmentation for patients with a primary anxiety disorder or a mood disorder: a pilot study. BMC Psychiatr. 2012;12:162. PMID: 23020711.

31.   Gillberg CH, R.; Steinhausen, H. A clinician's handbook of child and adolescent psychiatry. Cambridge: Cambridge University Press; 2006.

32.   Cohen JA, Bukstein O, Walter H, et al. Practice parameter for the assessment and treatment of children and adolescents with posttraumatic stress disorder. J Am Acad Child Adolesc Psychiatr. 2010 Apr;49(4):414-30. PMID: 20410735.

33.   Sareen J, Cox BJ, Goodwin RD, et al. Co-occurrence of posttraumatic stress disorder with positive psychotic symptoms in a nationally representative sample. J Traum Stress. 2005 Aug;18(4):313-22. PMID: 16281228.

34.   Kessler R, Berglund P, Demler O, et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Coborbidity Survey replication. Arch Gen Psychiatry. 2005;62:593-602. PMID: 15939837.

35.   Kilpatrick DG, Ruggiero KJ, Acierno R, et al. Violence and risk of PTSD, major depression, substance abuse/dependence, and comorbidity: results from the National Survey of Adolescents. J Consult Clin Psychol. 2003;71(4):692-700.

36.   Han C, Pae CU, Wang SM, et al. The potential role of atypical antipsychotics for the treatment of posttraumatic stress disorder. J Psychiatr Res. 2014 Sep;56:72-81. PMID: 24882700.

37.   Krystal JH, Rosenheck RA, Cramer JA, et al. Adjunctive risperidone treatment for antidepressant-resistant symptoms of chronic military service-related PTSD: a randomized trial. JAMA. 2011 Aug 3;306(5):493-502. PMID: 21813427.

38.   Cavanna AE, Servo S, Monaco F, et al. The behavioral spectrum of Gilles de la Tourette syndrome. J Neuropsych Clin N. 2009 Winter;21(1):13-23. PMID: 19359447.

39.   Tsertsvadze A, Maglione M, Chou R, et al. Updating comparative effectiveness reviews: current efforts in AHRQ's Effective Health Care program methods guide for effectiveness and comparative effectiveness reviews. Rockville, MD: Agency of Healthcare Research and Quality, 2008. PMID: 22091476. Available at: www.effectivehealthcare.ahrq.gov/

40.   Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. Agency of Healthcare Research and Quality Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 12-EHC047-EF. Rockville MD; Agency of Healthcare Research and Quality, 2012. PMID: 22479713. Available at: www.effectivehealthcare.ahrq.gov/

41.   Higgins JG, S. Chapter 9: Analyzing data and undertaking meta-analysis. Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration; 2009. Available at: www.handbook.cochrane.org

42.   De Hert M, Dobbelaere M, Sheridan EM, et al. Metabolic and endocrine adverse effects of second-generation antipsychotics in children and adolescents: a systematic review of randomized,

placebo controlled trials and guidelines for clinical practice. Eur Psychiat. 2011 Apr;26(3):144-58. PMID: 21295450.

43.   Maayan L, Correll CU. Weight gain and metabolic risks associated with antipsychotic medications in children and adolescents. J Child Adolesc Psychopharm. 2011 Dec;21(6):517-35. PMID: 22166172.

44.   Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. AHRQ; 2014. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/60/318/CER-Methods-Guide-140109.pdf.

45.   Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, et al. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol. 2000;53(9):964-72. [PMID: 11004423]

46.   Higgin JP, Green, S. Section 6.4.11.1 search filters. In The Cochrane Handbook for Systematic Reviews of Interventions The Cochrane Collaboration, 2011. Available at: www.handbook.cochrane.org

47.   Ip S, Hadar N, Keefe S, et al. A web-based archive of systematic review data. Systematic Reviews. 2012;1:15. PMID: 22588052.

48.   Higgins JP, Green, S. . Section 8. Assessing the risk of bias of included studies. In Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaroration; 2011. Available at: www.handbook.cochrane.org

49.   Berkman ND, Lohr KN, Ansari M, et al. Grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the Agency for Healthcare Research and Quality: An update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD; AHRQ: 2014. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/60/318/CER-Methods-Guide-140109.pdf.

50.   Wells G, Shea, B., O'connell, N., et al. The Newcastle-Ottawa Scale for assessing the quality of nonrandomized studies in meta-analyses. Department of Epidemiology and Community Medicine, University of Ottawa 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

51.   Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. BMJ. 2008 May 10;336(7652):1049-51. PMID: 18467413.

52.   Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105-24. PMID: 15449338.

53.   Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002 Aug 30;21(16):2313-24. PMID: 12210616.

54.   Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods. 2012 Jun;3(2):80-97. PMID: 26062083.

55.   Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med. 2014 Feb 18;160(4). PMID: 24727843.

56.   Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res. 2001 Aug;10(4):277-303. PMID: 11491414.

57.   Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ (Clinical research ed). 1997 Sep 13;315(7109):629-34. PMID: 9310563.

58.   Achana FA, Cooper NJ, Bujkiewicz S, et al. Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. BMC Med Res Methodol. 2014;14:92. PMID: 25047164.

59.   Wei Y, Higgins JP. Bayesian multivariate meta-analysis with multiple outcomes. Stat Med. 2013 Jul 30;32(17):2911-34. PMID: 23386217.

60. Achana FA, Cooper NJ, Dias S, et al. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. Stat Med. 2013 Feb 28;32(5):752-71. PMID: 22865748.
61. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64(12):1311-6. PMID: 21802902.
62. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence-- imprecision. J Clin Epidemiol. 2011 Dec;64(12):1283-93. PMID: 21839614.
63. Balshem H, Stevens A, Ansari M, et al. Finding grey literature evidence and assessing for outcome and analysis reporting biases when comparing medical interventions: AHRQ and the Effective Health Care Program Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville MD2008. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD; AHRQ: 2014. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/60/318/CER-Methods-Guide-140109.pdf

VIII. Definition of Terms

Not applicable.

IX. Summary of Protocol Amendments

In the event of protocol amendments, the date of each amendment will be accompanied by a description of the change and the rationale.

X. Review of Key Questions

For all EPC reviews, key questions are reviewed and refined as needed by the EPC with input from Key Informants and the TEP to assure that the questions are specific and explicit about what information is being reviewed. In addition, the key questions were posted for public comment and finalized by the EPC after review of the comments. This input is intended to ensure that the Key Questions are specific and relevant.

XI. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. Officers from AHRQ and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XII.   Technical Experts

Technical Experts comprise a multi-disciplinary group of clinical, content, and methodologic experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XIII.   Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer Reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XIV. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest which cumulatively total greater than $1,000 will usually disqualify EPC core team investigators.

XV.   Role of the Funder

This project was funded under Contract No. HHSA 290-2015-00001-I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

APPENDIX A: Summary tables of first and second generation antipsychotics included in the CER

Table A1. First generation antipsychotics included in the CER

Generic name	Trade names(s)	Mode of administration	Usual dose	Frequency
Chlorpromazine	Chlorpromazine hydrochloride Chlorpromazine hydrochloride	Oral	Adult, 200-600 mg/day; Children ≥6mo, 0.5 to 1mg/kg/dose	1-4 times q4-6 hr
	D/C : Intensol, Promapar, Sonazine, Thorazine	IM/IV	300-800 mg/day	q4-6 hr
Fluphenazine	Fluphenazine decanoate Fluphenazine hydrochloride	Oral	2.5-10 mg/day	3-4 times
		IM	2.5-10 mg/dose	Q6-8hr
	D/C: Fluphenazine, Permitil, Prolixin, Prolixin decanoate, Prolixin enanthate			
Haloperidol	Haloperidol Haldol	Oral Tablets	Adult, 4-12 mg/day	1-2 times
	Haloperidol decanoate	Solution	Children 3-12 yr, 0.5 to 0.15mg/kg/day	2-3 times
	D/C: Haldol solutab, Haloperidol intensol, Haloperidol lactate	IM (as lactate)	Adult, maximum 20mg/day	Every hr if needed
Loxapine	Loxapine, Loxapine succinate	Oral	Adult, 60-100 mg/day; Not recommended in children <16yr	2-4 times
	D/C : Loxitane, Loxitane C, Loxitane IM			
Molindone	Molindone hydrochloride	Oral	50-100 mg/day	3-4 times
	D/C : Moban			
Perphenazine	Perphenazine	Oral (non-hospitalized) Oral (hospitalized)	12-18 mg/day; 16-64 mg/day	3 times 2-4 times
	D/C : Trilafon			
Pimozide	Orap	Oral	Adults, 7-10 mg/day; Children initiate at 0.05 mg/kg once at bedtime; increase every third day to a maximum of 0.2 mg/kg not to exceed 10 mg/day	1-3 times

D/C = discontinued according to FDA site; IM = intramuscular; IV = intravenous

Table A1. First generation antipsychotics included in the CER (continued)

Generic name	Trade names(s)	Mode of administration	Usual dose	Frequency
Prochlorperazine	Compro Procomp	Oral	15-40 mg/day	3-4 times
	Prochlorperazine Prochlorperazine edisylate	IM	15-40 mg/day	3-4 times
	Prochlorperazine maleate	IV	7.5-40 mg/day	3-4 times
	D/C : Compazine	Rectal	25-50 mg/day	1-2 times
Thiothixene	Navane Thiothixene	Oral	6-30 mg/day	2-3 times
	D/C : Thiothixene hydrochloride, Thiothixene hydrochloride intensol			
Thioridazine	Thioridazine hydrochloride	Oral	150-300 mg/day	2-3 times
	D/C: Thioridazine hydrochloride intensol, Mellaril, Mellaril-S			
Trifluoperazine	Trifluoperazine hydrochloride	Oral (non-hospitalized)	Adult, 1-2 mg	2 times/day
	D/C: Stelazine	Oral (hospitalized)	Adult, 15-21 mg/day	2 times/day
			Children (6-12 yrs), 1 mg	1-2 times/day

D/C = discontinued according to FDA site; IM = intramuscular; IV = intravenous

Table A2. Second generation antipsychotics included in the CER

Generic name	Brand names(s)	Mode of administration	Recommended dose	Frequency
Aripiprazole	Aripiprazole Abilify Abilify Discmelt Abilify Maintena	Tablet Tablet Orally disintegrating tablet Extended-release injectable suspension	Adult schizophrenia, 10-15mg/day; Adolescent schizophrenia, 10mg/day; Adult BP (mania as monotherapy or adjunctive), 15mg/day Pediatric BD (mania as monotherapy of adjuctive), 10mg/day; Pediatric irritability with autistic disorder, 5-10 mg/day Tourette's disorder, ≥50kg 10mg/day; <50kg 5mg/day	QD Except for Abilify Maintena which is 400mg monthly
	Abilify	Injection	Adults - Agitation associated with schizophrenia or BD mania, 9.75 mg injected IM (max 30mg/day)	≥2 hr between doses
Asenapine	Saphris	Orally disintegrating tablet	Adult acute schizophrenia 5mg; Adult maintenance schizophrenia 5-10mg Adult BD mania (montherapy or adjunctive) 5-10mg; Child/adolescent BD mania 2.5-10 mg	2 times/day 2 times/day 2 times/day 2 times/day
Brexpiprazole	Rexulti	Tablet	Adult major depression 2 mg/day; Adult schizophrenia 2 to 4 mg/day	QD
Cariprazine	Vraylar	Capsule	Adult schizophrenia 1.5 mg to 6 mg/day Adult bipolar mania 3 mg to 6 mg/day	QD QD
Clozapine	Clozapine Clozaril	Tablet Orally disintegrating tablet	300-450 mg/day	1-3 times/day
Iloperidone	Fanapt	Tablet	Adult schizophrenia, 12-24mg/day	2 times/day
Lurasidone	Latuda	Tablet	Adult schizophrenia, 40-160mg/day Adults BD depression (monotherapy or adjunctive), 20-120mg/day	QD
Olanzapine	Olanzapine Zyprexa, Zyprexa Zydis Symbyax	Tablet Orally disintegrating tablet IM injection	Adult/Adolescent schizophrenia, 10mg/day; Adult BD I 10-15mg/day; Adolescent BD I 10mg/day	QD QD
Paliperidone	Invega Invega Sustenna	Tablet extended release IM injection	Adults, 3-12mg/day Adolescents, <51kg 3-6mg/day, >51kg 3-12 kg/day Adults, 39-234 mg (schizophrenia), 78-234 mg schizoaffective disorder	QD in the AM Monthly
Quetiapine	Quetiapine fumarate Seroquel Seroquel XR	Tablet Sustained release tablets	Adult schizophrenia, 150-750mg/day; Adolescent schizophrenia, 400-800mg/day; Adult BD (mania), 400-800mg/day; Child/adolescent BD (mania), 400-600mg/day; Adult BD (depression), 300mg/day; Adult BD (maintenance), 400-800mg/ Adult (depression with antidepressants), 150-300mg/day	2 times/day 2 times/day 2 times/day 2 times/day (Seroquel XR, QD at

Generic name	Brand names(s)	Mode of administration	Recommended dose	Frequency
				bedtime)
Risperidone	Risperidone, Risperdal, Risperdal consta	Tablet Solution Orally disintegrating tablet IM injection	Adult schizophrenia, 4-8mg/day; Adolescent schizophrenia, 3mg/day; Adult BD (mania), 1-6mg/day; Child/adolescent BD (mania), 1-2.5mg/day; Irritability childhood autism (≥ 5yrs), 0.5mg/day (<20kg); 1mg/day (≥ 20 kg)	1-2 times/day
Ziprasidone	Ziprasidone hydrochloride Geodon	Capsules Oral suspension IM injection	Adult schizophrenia, up to 80mg; BD (manic/mixed, maintenance), 40-80mg; Agitation associated with schizophrenia (IM), up to max 40mg/day	2 times/day 2 times/day 10mg may be injected q 2 hr

BD = bipolar disorder; IM = intramuscular; QD = every day

Table A3. First generation antipsychotics: FDA status

Drug	FDA status	Indications	Age group approved for	Black box Warnings
Chlorpromazine	Approved 1974	Schizophrenia BP (mania) Hyperactivity Uncontrolled hiccups, nausea and vomiting	Adults Children (1-12 years)	Patients with cardiovascular disease or hx of seizures
Droperidol	Approved 1970			Cases of QT prolongation and/or torsade de pointes have been reported in patients receiving droperidol at doses at or below recommended doses.
Fluphenazine	Approved 1960	Schizophrenia BD (mania)	Adults Children >12yrs Not recommended for use in children under 12 years	Possible increased mortality in elderly with dementia-related psychosis Not approved for the treatment of dementia-related behavior problems.
Haloperidol	Approved 1986	Schizophrenia Tourette's Disorder	Adults Safety and effectiveness in pediatric patients <18 years have not been established	Increased mortality in elderly with dementia-related psychosis
Loxapine	Approved 1975	Schizophrenia	Adults Safety and effectiveness in pediatric patients <16 have not been established	Increased mortality in elderly with dementia-related psychosis
Molindone	Approved 1974	Schizophrenia	Adults Use in pediatric patients <12 years is not recommended because safe	Increased mortality in elderly with dementia-related psychosis

Drug	FDA status	Indications	Age group approved for	Black box Warnings
			and effective conditions for its usage have not been established	
Perphenazine	Approved 1965	Schizophrenia	Adults Safety and effectiveness in pediatric patients have not been established	Increased mortality in elderly with dementia-related psychosis Children, adolescents, and young adults taking antidepressants are at increased risk of suicidal thinking
Pimozide	Approved 1984	Tourette's Disorder	Children and adults 8-53 years. Limited evidence in children <12 years Use and safety have not been evaluated in other childhood disorders, ORAP is not recommended for use in any condition other than Tourette's Disorder	Use of pimozide in treatment of Tourette's Disorder involves different risk/benefit considerations than treatment of other conditions. Tardive Dyskinesia Neuroleptic Malignant Syndrome (NMS) Sudden, unexpected deaths in conditions other than Tourette's Disorder. May have tumorigenic potential.
Prochlorperazine	Approved 1956	Schizophrenia Severe nausea and vomiting	Adults and children Children >2 years and > 20 pounds	May cause tardive dyskinesia
Thiothixene	Approved 1967	Schizophrenia	Adults Safety and effectiveness in pediatric patients <12 years have not been established	Increased mortality in elderly with dementia-related psychosis
Thioridazine	Approved 1962	Schizophrenia	Adults and children	Life-threatening pro-arrhythmic effect
Trifluoperazine	Approved 1959	Schizophrenia	Adults and children (6-12 yrs)	Increased mortality in elderly patients with dementia-related psychosis

BP = bipolar disorder

Table A4. Second generation antipsychotics: FDA status

Drug	FDA status (year approved)	Indications	Age group approved for	Black box warnings
Aripiprazole	2002	Schizophrenia	Adults & adolescents (13-17 yr)	Increased mortality in elderly with dementia-related psychosis
	2004	BD (manic/mixed) monotherapy or adjunctive to lithium or valproate	Adults & pediatrics (10 -17 yr)	Children, adolescents, and young adults taking antidepressants are at increased risk of suicidal thinking & behaviour
	2007	Adjunctive tx of major depressive disorder Adults with agitation associated with schizophrenia or BD(L) (manic/mixed)	Adults	Leukopenia, Neutropenia, Agranulocytosis Not approved for behavior problems in older adults with dementia-related psychosis.
	2009	Irritability in autistic disorder, injection	Children (6-17 yr)	
	2014	Tourette's disorder	Patients 6+ yr	
Asenapine	2009	Acute schizophreniaBD I (manic/mixed)	Adults Adults	Increased mortality in elderly with dementia-related psychosis
	2010	Maintenance of schizophrenia BD (manic/mixed) adjunctive to lithium or valproate	Adults Adults	
	2015	BD (manic)	Children (10-17 yr)	
Brexpiprazole	2015	Adjunctive tx of major depressive disorder Schizophrenia	Adults Adults	Increased mortality in elderly with dementia-related psychosis Suicidal thoughts and behaviors
Cariprazine	2015	Schizophrenia Bipolar mania	Adults Adults	Increased mortality in elderly with dementia-related psychosis
Clozapine	1989	Treatment resistant schizophrenia	Adults Pediatric use: safety & effectiveness not established in patients <18 yr	1. agranulocytosis 2. seizures 3. myocarditis
	2002	Reduce the risk of suicidal behavior in younger schizophrenics.		4. cardiovascular and respiratory effects, (respiratory and/or cardiac arrest). 5. increased mortality in elderly patients with dementia-related psychosis
Iloperidone	2009	Acute schizophrenia	Adults	Increased mortality in elderly with dementia-

Drug	FDA status (year approved)	Indications	Age group approved for	Black box warnings
				related psychosis Not approved for patients with dementia-related psychosis.
Lurasidone	2010	Schizophrenia, 40-160mg/day	Adults	Elderly patients with dementia-related psychosis treated with antipsychotic drugs are at an increased risk of death.
	2013	BD (depressive) monotherapy or adjunctive	Adults	Not approved for the treatment of patients with dementia- related psychosis Increased risk of suicidal thinking and behavior in children, adolescents, and young adults taking antidepressants Monitor for worsening and emergence of suicidal thoughts and behaviors
Olanzapine	1996	Schizophrenia & BD(L) (manic/mixed)	Adults Adolescents (13-17 yr), schizophrenia & BD (manic/ mixed)	Increased mortality in elderly with dementia-related psychosis Not approved for patients with dementia-related psychosis. Increased risk of suicidal thinking and behavior in children, adolescents, and young adults taking antidepressants
	2003: combined w fluoxetine	BD (depressive)		
	2004	BD(L) long-term tx		
	2009: combined w fluoxetine	Tx resistant depression		
	2013: combined w fluoxetine		Pediatric use (10-17 yr) BD (depression)	
Paliperidone	2006	Schizophrenia Schizoaffective disorder	Adult	Increased mortality in elderly with dementia-related psychosis
	2011	Schizophrenia	Adolescents (12-17)	
			Pediatric use: safety & effectiveness not established in patients <18 yr	
Quetiapine	1997	Schizophrenia	Adults & adolescents (13-17 yr) Adults, children & adolescents (10-17 yr)	Increased mortality in elderly with dementia-related psychosis Increased risk of suicidal thoughts and behaviors in children, adolescents and young adults taking antidepressants Not approved for patients with dementia-related psychosis
	2004	BD (acute manic; monotherapy or adjunct to lithium or divalproex))	Adults Adults	
	2008	BD (depression)		

Drug	FDA status (year approved)	Indications	Age group approved for	Black box warnings
		BD (maintenance)		
	2013	BD (mania; monotherapy)	Children & adolescents (10-17)	
Risperidone	1993	Schizophrenia	Adults & adolescents (13-17 yr)	Increased mortality in elderly with dementia-related psychosis
	2003	BD (manic/mixed)	Adults & adolescents (10-17 years)	
	2007	Irritability associated with autism	Children (5-16 yr)	
Ziprasidone	2001	Schizophrenia	Adults	Increased mortality in elderly with dementia-related psychosis
	2004	BD (manic/mixed)	Adults	
	2009	BD (maintenance)	Adults	
	2014	Agitation in schizophrenia	Adults	
			Pediatric use: safety & effectiveness not established in patients <18 yr	

BD= bipolar disorder

Appendix B: MEDLINE search terms and strategy

1. Adjustment Disorders/
2. Anorexia/
3. Anxiety/
4. exp Anxiety Disorders/
5. exp "Attention Deficit and Disruptive Behavior Disorders"/
6. exp Behavioral Symptoms/
7. Child Behavior Disorders/
8. exp Child Development Disorders, Pervasive/
9. exp Eating Disorders/
10. exp Hyperphagia/
11. exp Impulse Control Disorders/
12. exp Impulsive Behavior/
13. Irritable Mood/
14. Mental Disorders/
15. exp Mood Disorders/
16. Movement Disorders/
17. "Off-Label Use"/
18. Psychomotor Agitation/
19. Rett Syndrome/
20. exp "Schizophrenia and Disorders with Psychotic Features"/
21. Schizophrenia, Childhood/
22. exp Sleep Disorders/
23. exp Substance-Related Disorders/
24. exp Tic Disorders/
25. Violence/
26. (ADHD* or (attention deficit adj2 disorder*) or hyperkinetic syndrome).tw,kf.
27. ((adjustment or reactive) adj disorder*).tw,kf.
28. (affective adj2 (disorder* or disregulation or dysregulation)).tw,kf.
29. (aggressi* or agitat*).tw,kf.
30. agoraphobi*.tw,kf.
31. ((alcohol* or drug* or cannabi* or cocaine* or heroin or marijuana* or narcotic* or opiate* or opioid* or substance*) adj2 (abus* or addict* or depend* or disorder* or withdrawal*)).tw,kf.
32. ((addicti* or compulsi* or explosive or impuls*) adj2 (behavio* or disorder*)).tw,kf.
33. (((anankastic or compulsiv* or obsessive) adj (behavio* or disorder* or neuros* or personalit*)) or OCD).tw,kf.
34. anorexi*.tw,kf.
35. anxiety.tw,kf.
36. (autis* or asperger* or kanner* syndrome).tw,kf.
37. (behavio* adj2 (disorder* or disturb* or disrupt* or dyscontrol* or illness* or issue* or outburst* or problem*)).tw,kf.
38. (((behavio* or disorder* or episod*) adj (hypomanic or manic)) or mania*).tw,kf.
39. (binge adj (drink* or eat*)).tw,kf.
40. (bi polar or bipolar).tw,kf.
41. bulimi*.tw,kf.

42. (claustrophobi* or phobia* or phobic).tw,kf.
43. ((combat or war) adj (disorder* or neuros*)).tw,kf.
44. conduct disorder*.tw,kf.
45. cyclothymi*.tw,kf.
46. ((defiant or disrupt* or oppositional) adj (behavio* or disorder*)).tw,kf.
47. delusion*.tw,kf.
48. dementia praecox.tw,kf.
49. depress*.tw,kf.
50. ((dis integrative or disintegrative or dys integrative or dysintegrative) adj disorder*).tw,kf.
51. (dys somnia* or dyssomnia* or insomnia* or para somnia* or parasomnia*).tw,kf.
52. dysthymi*.tw,kf.
53. eating disorder*.tw,kf.
54. ((emotion* or mood) adj2 (disorder* or dis regulation or disregulation or dys regulation or dysregulation)).tw,kf.
55. (hoarder* or hoarding).tw,kf.
56. (hyper activ* or hyperactiv*).tw,kf.
57. hyperphagia*.tw,kf.
58. irritab*.tw,kf.
59. kleptomania*.tw,kf.
60. (minimal brain adj (dis function* or disfunction* or dys function* or dysfunction*)).tw,kf.
61. (mood adj2 (labil* or swing*)).tw,kf.
62. (off label* or offlabel* or unlabeled indication* or unlabeled use*).tw,kf.
63. (panic* adj (attack* or disorder*)).tw,kf.
64. (para suicid* or parasuicid*).tw,kf.
65. paranoi*.tw,kf.
66. pervasive development* disorder*.tw,kf.
67. ((post traumatic or posttraumatic) adj2 (disorder* or neuros*)).tw,kf.
68. ((psycho* or sociopath*) adj (disorder* or personalit*)).tw,kf.
69. psychos*.tw,kf.
70. PTSD*.tw,kf.
71. (rett* adj (syndrome* or disorder*)).tw,kf.
72. (self adj (destruct* or harm* or injur* or mutilat*)).tw,kf.
73. (schizo affect* or schizoaffect*).tw,kf.
74. schizophreni*.tw,kf.
75. shell shock*.tw,kf.
76. (sleep adj2 (disorder* or dysfunction*)).tw,kf.
77. stress disorder*.tw,kf.
78. tourette*.tw,kf.
79. tic disorder*.tw,kf.
80. unstable mood*.tw,kf.
81. violen*.tw,kf.
82. or/1-81
83. exp Antipsychotic Agents/
84. exp Butyrophenones/

85. exp Phenothiazines/
86. exp Thioxanthenes/
87. abilify.mp.
88. adasuve.mp.
89. aldazine.mp.
90. anatensol.mp.
91. anti naus.mp.
92. (anti psychotic* or antipsychotic*).mp.
93. aripiprazole.mp.
94. 82VFR53I78.rn.
95. arizole.mp.
96. asenapine.mp.
97. JKZ19V908O.rn.
98. atrolak.mp.
99. biquelle.mp.
100. brexpiprazole.mp.
101. 2J3YBM1K8C.rn.
102. buccastem.mp.
103. calmazine.mp.
104. cariprazine.mp.
105. chloractil.mp.
106. chlorpromanyl.mp.
107. chlorpromazine.mp.
108. U42B7VYA4P.rn.
109. clopine.mp.
110. clozapine.mp.
111. J60AR2IKIC.rn.
112. clozaril.mp.
113. compazine.mp.
114. compro.mp.
115. decazate.mp.
116. delucon.mp.
117. denzapine.mp.
118. dozic.mp.
119. droleptan.mp.
120. droperidol.mp.
121. O9U0F09D5X.rn.
122. ebesque.mp.
123. fanapt.mp.
124. fazaclo.mp.
125. fazalco.mp.
126. fentazin.mp.
127. fluphenazine.mp.
128. S79426A41Z.rn.
129. fortunan.mp.
130. geodon.mp.

131. haldol.mp.
132. halo peridol.mp.
133. haloperidol.mp.
134. J6292F8L3D.rn.
135. halperon.mp.
136. iloperidone.mp.
137. 133454-47-4.rn.
138. inapsine.mp.
139. invega.mp.
140. lanzek.mp.
141. largactil.mp.
142. latuda.mp.
143. loxapac.mp.
144. loxapine.mp.
145. LER583670J.rn.
146. loxitane.mp.
147. lurasidone.mp.
148. 22IC88528T.rn.
149. (major adj (tranquili?er* or tranquilli?er*)).mp.
150. mellaril*.mp.
151. melleril.mp.
152. mintreleq.mp.
153. moban.mp.
154. modecate.mp.
155. moditen.mp.
156. molindone.mp.
157. RT3Y3QMF8N.rn.
158. nausetil.mp.
159. navane.mp.
160. neuroleptic*.mp.
161. novo flurazine.mp.
162. novo peridol.mp.
163. novo ridazine.mp.
164. novo trifluzine.mp.
165. nu prochlor.mp.
166. olanzaccord.mp.
167. olanzapine.mp.
168. 132539-06-1.rn.
169. orap.mp.
170. ormazine.mp.
171. ozidal.mp.
172. ozin.mp.
173. paliperidone.mp.
174. 838F01T721.rn.
175. permitil.mp.
176. perphenazine.mp.

177. FTA7XXY4EZ.rn.
178. pimozide.mp.
179. 1HIZ4DL86F.rn.
180. procalm.mp.
181. prochlorazine.mp.
182. prochlorperazine.mp.
183. YHP6YLT61T.rn.
184. procomp.mp.
185. prolixin.mp.
186. promapar.mp.
187. prorazin.mp.
188. protran.mp.
189. proziere.mp.
190. prozine.mp.
191. quetiapine.mp.
192. BGL0JSY5SI.rn.
193. quetiaccord.mp.
194. quetin.mp.
195. resdone.mp.
196. rexulti.mp.
197. rideril.mp.
198. rispa.mp.
199. risperdal.mp.
200. risperidone.mp.
201. L6UH7ZF8HC.rn.
202. rispernia.mp.
203. rixadone.mp.
204. saphris.mp.
205. seotiapim.mp.
206. sequase.mp.
207. serenace.mp.
208. seronia.mp.
209. seroquel.mp.
210. solazine.mp.
211. sonazine.mp.
212. sondate.mp.
213. stelazine.mp.
214. stemetil.mp.
215. stemzine.mp.
216. sycrest.mp.
217. syquet.mp.
218. terfluzine.mp.
219. thioridazine.mp.
220. N3D6TG58NI.rn.
221. thiothixene.mp.
222. 7318FJ13YJ.rn.

223. thorazine.mp.
224. tiotixene.mp.
225. trifluoperazine.mp.
226. 214IZI85K3.rn.
227. trilafon.mp.
228. versacloz.mp.
229. vertigon.mp.
230. vraylar.mp.
231. xeplion.mp.
232. xomolix.mp.
233. xylac.mp.
234. zaluron.mp.
235. zaponex.mp.
236. zeldox.mp.
237. ziprasidone.mp.
238. 6UKA5VEJ6X.rn.
239. zylap.mp.
240. zypadhera.mp.
241. zypine.mp.
242. zyprexa.mp.
243. or/83-242
244. and/82,243
245. Adolescent/
246. Adolescent Medicine/
247. exp Child/
248. exp Minors/
249. exp Pediatrics/
250. exp Puberty/
251. Students/
252. Young Adult/
253. adolescen*.mp.
254. (boy* or girl* or teen*).mp.
255. (child* or grade school* or kid or kids or kindergar?en* or minors* or preschool* or pre school* or school age* or schoolchild* or toddler*).mp.
256. ((colleg* or high school* or highschool* or middle school* or universit*) adj2 (age* or student*)).mp.
257. (paediatric* or peadiatric* or pediatric*).mp.
258. (prepubescen* or pubescen* or pubert*).mp.
259. (young* adj (adult* or men or mens or people* or person* or women*)).mp.
260. (youth or youths).mp.
261. or/245-260
262. and/244,261
263. exp Epidemiologic Studies/
264. controlled clinical trial.pt.
265. randomized controlled trial.pt.
266. drug therapy.fs.

267. (case control or cohort* or follow up or followup or longitudinal or prospective* or retrospective).tw,kf.
268. ((compari* or epidemiologic* or experimental or observational) adj2 (analy* or study or studies)).tw,kf.
269. groups.ab.
270. placebo.ab.
271. random*.ab.
272. trial.ab.
273. or/263-272
274. exp animals/ not humans.sh.
275. 273 not 274
276. and/262,275
277. (case reports or comment or editorial or letter).pt.
278. 276 not 277
279. limit 278 to english
280. limit 279 to yr="1987-current"

Appendix C: Preferred symptom outcome measurement tools from original CER

Outcome measurement tool*
Aberrant Behavior Checklist
Behavior Problems Inventory
Brief Psychiatric Rating Scale
Childhood Autism Rating Scale
Child Behavior Checklist
Child Mania Rating Scale
Children's Aggression Scale
Children's Depression Rating Scale
Children's Psychiatric Rating Scale
Children's Yale-Brown Obsessive-Compulsive Scale
Conners Parent/Teacher Rating Scale
Gilliam Autism Rating Scale
General Behavior Inventory
Hamilton Anxiety Rating Scale
Hamilton Depression Rating Scale
Nisonger Child Behavior Rating Scale
Overt Aggression Scale
Rating of Aggression Against People and/or Property Scale
Personal Assessment Checklist
Positive and Negative Symptom Scale
Ritvo-Freeman Real Life Rating Scale
Scale for the Assessment of Negative Symptoms
Scale for the Assessment of Positive Symptoms
Social and Occupational Functioning Assessment Scale
Strength and Difficulties Questionnaire
Tic Symptom Self-Report
Total Child Symptom Inventory
Vineland Adaptive Behavior Scales
Yale Global Tics Severity Score
Young Mania Rating Scale

*we will modify this list based on validated tools used in new studies, especially for those reporting on a condition added to this review