Comparative Effectiveness Review
Number 184

Effective Health Care Program

First- and Second-Generation Antipsychotics in Children and Young Adults: Systematic Review Update

Executive Summary

Introduction

The use of psychotropic medications, including antipsychotics, in children, adolescents, and young adults has risen over the past 20 years,[1-6] and use of antipsychotics in children with public health insurance[2] and living in foster homes[4] is greater than for those with private health insurance in the United States. During 2010, the percentages of young people filling prescriptions for antipsychotics in the United States was 0.11 percent (younger children), 0.8 percent (older children) 1.19 percent (adolescents), and 0.84 percent (young adults).[5] Antipsychotic medications are commonly categorized into two classes. First-generation antipsychotics (FGAs) were developed in the 1950s, while second-generation antipsychotics (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

Purpose of Review:

To assess the effectiveness and harms of first- and second-generation antipsychotics (FGAs and SGAs) used for treating children, adolescents, and young adults.

Key Messages:

- There was little information directly comparing different antipsychotics, on patient-important outcomes including quality of life, and on outcomes for young children.
- FGAs probably cause less weight gain than SGAs, and (for schizophrenia) there may be little or no difference between the classes for reducing symptoms and illness severity.
- SGAs probably improve to some extent symptoms for which they are usually prescribed, but also cause adverse effects including weight gain, high triglyceride levels, extrapyramidal symptoms, sedation, and somnolence.
- More research is needed comparing the effects of different antipsychotics over the long-term and developing monitoring systems.

AHRQ Agency for Healthcare Research and Quality
Effective Health Care Program

over the last two decades.[7] A shift towards SGAs was partly driven by the lower risk of extrapyramidal symptoms with their use, 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 seems more prone than FGAs to adverse effects such as weight gain, elevated lipid and prolactin levels, and development of metabolic syndrome.[8-10] 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.[11] 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 use of antipsychotics is common in children and adults.[1,12] Twenty-four to 31 percent of antipsychotic-treated children have attention deficit hyperactivity disorder (ADHD),[1,13] and 34.5 percent of antipsychotic-treated young adults have depression.[5] In Medicaid-enrolled children, ADHD accounted for 50 percent of total antipsychotic use in 2007;[12] ADHD and mood disorders not otherwise specified were the most common uses (32% and 37.2%, respectively) for antipsychotics in a sample of Medicaid-insured children in Vermont during 2012.[12] In these cases or other conditions such as conduct disorders, antipsychotics are usually given for adjunctive treatment of severe behavioral symptoms (e.g., aggression), rather than for psychoses.[5-14] 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;[12] these uses are accompanied by considerable controversy because of concerns regarding the balance of benefits and harms. This is particularly relevant when other treatment options exist for many conditions; for instance, fewer than half of very young, privately insured children taking antipsychotics received formal mental health services in 2007.[1]

Because of the marked increase in FDA-approved 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),[15] 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,[16] but these parameters may be considered outdated (all studies cited in the parameters were published prior to 2012) for providing the best evidence. The purpose of the systematic review is to provide a comprehensive synthesis of the evidence examining the benefits and harms associated with the use of FDA-approved FGAs and SGAs in children, adolescents, and young adults ≤24 years of age. This systematic review 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.

This is an update of Comparative Effectiveness Review (CER) No. 39 published in 2012.[17] The scope of this update has remained quite similar, with key changes being the addition of (1) three newly approved SGAs (i.e., brexpiprazole, asenapine, lurasidone) and the previously discontinued FGA molindone, (2) some conditions of interest (i.e., anxiety, depression, substance use), and (3) modification to some key outcomes to be more specific to symptoms targeted by clinicians when prescribing antipsychotics.

Scope of Review and Key Questions

Conditions of Interest

- Schizophrenia and schizophrenia-related psychoses, including schizoaffective disorder and 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
- 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's syndrome).
- Behavioral issues outside the context of a mental disorder, including aggression, agitation, behavioral dyscontrol, irritability, self-injurious behaviors, and insomnia.

Key Questions

Key Question 1. For each condition of interest, 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?

Key Question 2. 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)?

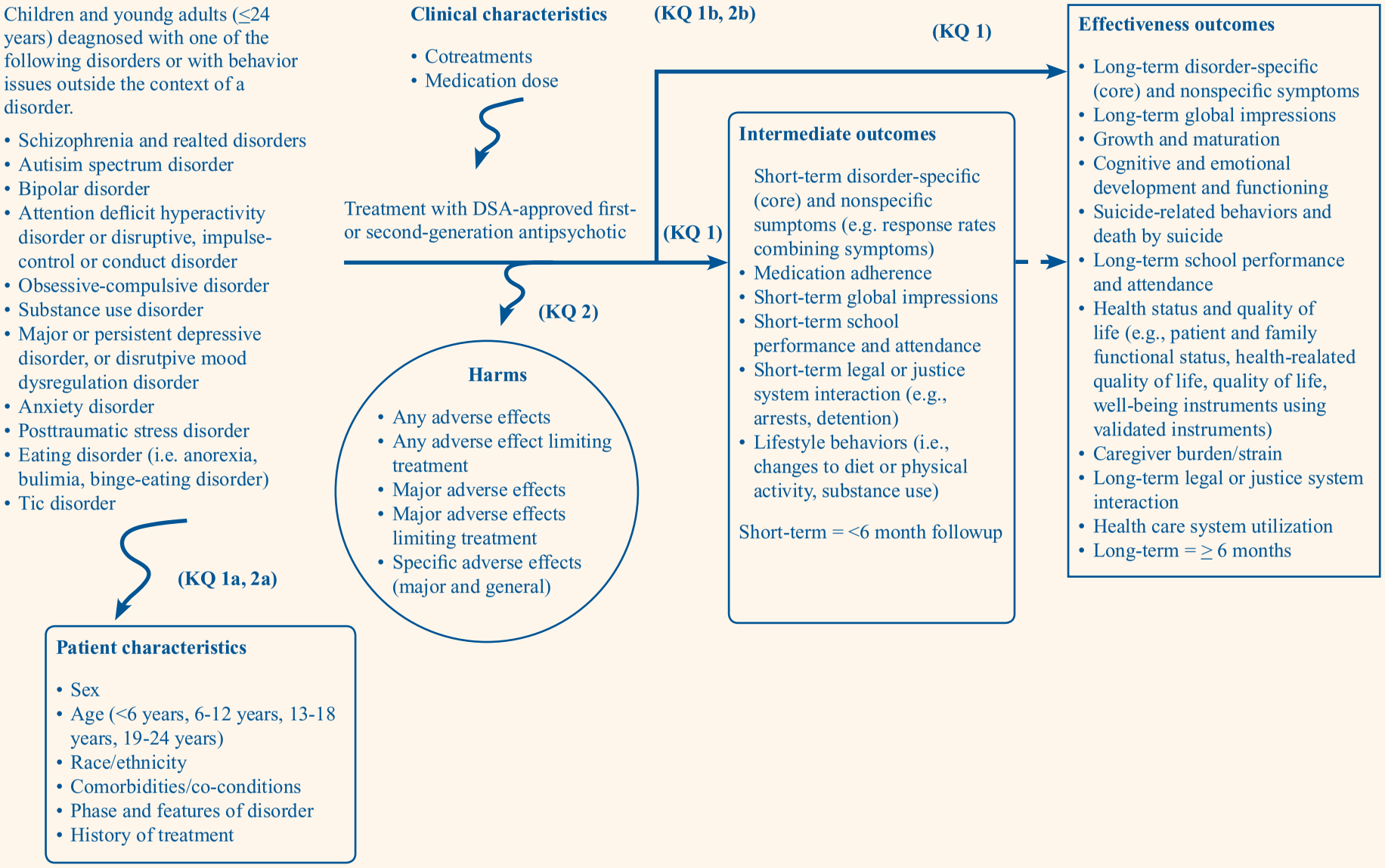
a. Do the harms vary with respect to patient characteristics, such as age, sex, race/ethnicity, diagnosis, medical comorbidities, phase of disorder, and prior exposure to antipsychotics?

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?

Analytic Framework

Figure A is an analytic framework that depicts the structure used to address the Key Questions (KQs) for evaluating the benefits and harms of FGAs and SGAs in children and young adults (≤24 years of age). We examined 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 was determined (by condition) for intermediate outcomes (e.g., disorder-specific and nonspecific symptoms, medication adherence, and lifestyle behaviors from short-term treatment durations), and effectiveness outcomes (e.g., symptoms over long-term treatment, growth and maturation, health status and quality of life, caregiver burden/strain). In KQ2, we assessed harms across conditions in terms of adverse effects (AEs) categorized as major (e.g., mortality, development of diabetes) and general (e.g., extrapyramidal effects, weight gain, hyperprolactinemia). Within each KQ, we assessed outcomes for subgroups of patients or studies based on patient and clinical/treatment characteristics.

Figure A. Analytic framework for the Key Questions evaluating the comparative effectiveness of FDA-approved first- and second generation antipsychotics in children and young adults 24 years old and under



FDA = Food and Drug Administration; KQ = Key Question

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).[18] We provide here a summary of the methods outlined in detail in the protocol and full report.[19]

Inclusion/Exclusion Criteria

We used the eligibility criteria in terms of the population, intervention(s), comparator(s), outcome(s), timing (of followup), setting, and design of study (PICOTS-D) as presented in Table A; details specific to our key outcomes follow. The primary focus in KQ2 was harms across all conditions, rather than within each condition, because adverse events associated with an antipsychotic are likely to be consistent regardless of the indication for which a drug is being taken; the difference in harms between conditions was treated as a subgroup of interest. We defined nonrandomized controlled trials (NRCTs) as experimental trials without random allocation but where intervention(s) are introduced, standardized, and allocated objectively [e.g., by date of birth, but not using subjective means such as patient or clinician preferences] by investigators and blinding of participants is typically possible.

Table A. 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/DICD, ASD, BD, DD, ED, OCD, PTSD, SUD, SZ, TD, or behavioral issues outside the context of a disorder (e.g., insomnia). KQ1: For each condition category, inclusion of studies enrolling ≥90 percent of patients diagnosed with the specific condition (s). KQ2: Across all conditions, inclusion of studies enrolling patients within a single or within multiple/mixed condition categories. Subpopulations based on patient characteristics: sex; age; race/ethnicity; comorbidities/co-conditions; history of treatment; phase and features of disorder.
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) All formulations and doses eligible. 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	KQ1: intermediate and effectiveness outcomes (see following list of outcomes). KQ2: any 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; see following list of outcomes)
Timing	No minimum followup duration Short term: <6 months Long term: ≥6 months-<12 months; 12 months+
Setting	Any setting
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 disorders; ADHD/DICD = attention-deficit/hyperactivity disorder, or disruptive, impulse-control, or conduct disorders; AE = adverse effect; ASD = autism spectrum disorders; BD = bipolar disorder; DD = depressive disorders, 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 and related psychosis; TD = tic disorders

Outcomes

The key intermediate and effectiveness outcomes of interest to this review are listed below, followed by the harms. We accounted for duration of response, that is, short- (< 6 months) and long-term (≥ 6 months - < 12 months; ≥ 12 months).

Key Intermediate Outcomes

- Short-term (in terms of followup) disorder-specific (core) symptoms:
  - Schizophrenia and related psychoses: positive and negative symptoms;
  - Autism spectrum disorders: irritability, qualitative impairment in social interactions, communication, restricted repetitive and stereotyped behaviors;
  - Bipolar disorder: severity of mania, depression, psychotic features;
  - Attention deficit hyperactivity disorder or disruptive, impulse-control, and conduct disorders: aggression, 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;
  - Anxiety disorder: anxiety, irritability;
  - Posttraumatic stress disorder: hyperarousal, avoidance behaviors, intrusion;
  - Eating disorders: weight, 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, irritability, mood lability, self-injurious behaviors, and sleep latency and duration.
- Short-term nonspecific or associated symptoms
  - Response rates (other symptoms as reported were included but not considered key outcomes)
- Short-term global impressions and functioning

Key Effectiveness (Patient- and Family-Important) Outcomes

- Long-term disorder-specific symptoms (see list above)
- Long-term nonspecific or associated symptoms (see above)
- Long-term (≥ 6 month followup) global impressions and functioning
- Cognitive and emotional development and functioning
- Suicide-related ideations or behaviors, or death by suicide
- Generic and specific health status and quality of life (including patient and family functional status, well-being) using validated instruments
- Long-term (≥ 6 month followup) legal or justice system interaction

Key Harms: Major Adverse Effects

- Mortality, cerebrovascular disease-related events, development of diabetes mellitus, diabetic ketoacidosis, neuroleptic malignant syndrome, seizures, tardive dyskinesia, cardiomyopathies, cardiac arrhythmias, agranulocytosis and related (e.g., neutropenia)

Key Harms: General Adverse Effects

- Neuromotor effects: extrapyramidal symptoms including dystonia, akinesia, akathisia
- Metabolic effects: metabolic syndrome, change in body composition, fasting glucose, insulin sensitivity/ resistance, dyslipidemia, 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)
- Somnolence

Literature Search Strategy

We comprehensively searched the following electronic databases: Ovid MEDLINE and Ovid MEDLINE® In-Process & Other Non-Indexed Citations (1946 to Present), Cochrane Central Register of Controlled Trials via Wiley Cochrane Library (1991 to Present), EMBASE® via Ovid (1980 to 2016 Week 15), CINAHL Plus with Full Text via EBSCOhost (1937 to Present), PsycINFO® via Ovid (1987 to October Week 1, 2016), ProQuest® Dissertations and Theses Global (1861 to Present), and TOXLINE via The U.S. National Library of Medicine (1840s to Present). The

original searches from October 2015 were updated in April 2016. Several other sources were used to obtain studies or additional data, including reference lists of relevant systematic reviews and guidelines, ClinicalTrials.gov, and World Health Organization's International Clinical Trials Registry Platform. Drug manufacturers and other relevant stakeholders were notified of the opportunity to submit scientific information relevant to the interventions of this systematic review. We handsearched the Journal of Child and Adolescent Psychopharmacology, and the Journal of the American Academy of Child and Adolescent Psychiatry (2014-2015). We searched Drugs@FDA for Medical/Clinical and Statistical review documents containing harm data for patients 18 years of age or younger.

Study Selection

For the database searches, two reviewers independently screened the titles and abstracts (when available) using broad inclusion/exclusion criteria. One reviewer conducted all other searches outlined in the above section. Disagreements on final inclusion of all studies were resolved through consensus or third party adjudication.

Data Abstraction

One review team member extracted data for each study, and a senior level team member verified all data. A wide variety of checklists and scales were used to assess symptomatology in patients. In various instances (e.g., hyperactivity, aggression) we used subscale items on one or more questionnaires, rather than their overall composite scores, to capture the outcomes of interest with more specificity. Data on within-study subgroup analysis was collected.

Assessment of Methodological Quality of Individual Studies

Two experienced reviewers independently assessed the methodological quality of all original and new studies and resolved discrepancies through consensus. We re-assessed original studies because of changes to guidance in the Evidence-based Practice Center (EPC) program made subsequent to the original review. For randomized controlled trials (RCTs) and NRCTs we used the Cochrane Collaboration Risk of Bias tool,[20] with some modification based on EPC Methods guidance.[18] For cohort studies, we used the Newcastle-Ottawa Quality Assessment Scale.[21] Ratings reflect risk of bias (ROB: high, medium, low) such that the methodological quality is opposing (e.g., high ROB represents low quality).

Data Synthesis

For each KQ, we synthesized data in the following order based on type of drug comparison (as possible depending on data): aggregate (across class) data for FGAs versus SGAs, individual FGAs versus SGAs, within-class comparisons between individual FGAs and individual SGAs (other drug or dose), and then individual and aggregate data for FGAs versus placebo/no treatment and SGAs versus placebo/no treatment.

For pairwise meta-analyses, we employed a Bayesian random effects model.[22, 23] We used this approach when more than two studies reported on the same outcome and comparison. When different outcomes were considered to measure the same construct (e.g., different subscores of hyperactivity) we combined the results (at followup) of multiple scores using a standardized mean difference (SMD); in this way we were able to use as many studies as possible to capture effect estimates for our outcomes. When the SMD was not used because of reporting by multiple studies using the same measurement scale (enabling calculation of a mean difference [MD]), change scores were preferred over followup scores and we combined these two when necessary. We report MDs, SMDs, or risk ratios (RRs) with corresponding 95% credible intervals (95% CrI; Bayesian approaches provide variances using credible rather than confidence intervals, interpretable as the range of values within which there is a 95% chance of finding the true value of the effect). We often started with combining all studies within a condition category and then used our a priori defined list of patient and intervention subgroups (listed in Figure A as patient and clinical characteristics) to explore the heterogeneity. For intermediate and effectiveness outcomes we considered combining results from RCTs with NRCTS, but not with cohort studies. For harm outcomes we combined data from all study designs for the following reasons: 1) empirical evidence has found no difference in estimates of harms between meta-analyses of RCT and cohort study designs;[24] 2) a major contributor to bias on harms from observational studies is confounding by indication (e.g., differential prescriptions based on beliefs/knowledge about factors related to development of harms) which we did not believe was an important threat in studies examining mostly unanticipated harms in treatment naïve children; and 3) cohort studies are commonly recognized as contributing valuable, relatively high-quality evidence on harms applicable to real-world settings. To avoid making conclusions from these analyses without carefully

considering possible biases, we identified important potential confounders on which to assess the findings for heterogeneity and also extracted data from all studies that reported within-study subgroup analysis for possible patient and clinical treatment modifiers. In the event that results from studies were not combined, the findings of each study are reported with statistical precision indicated with confidence intervals (95% CIs).

For commonly reported key harm outcomes (weight and body mass index [BMI]), we employed a network meta-analysis to simultaneously evaluate a suite of comparisons including indirect comparisons (e.g., incorporation of placebo/no treatment-controlled and head-to-head trial data) while still preserving the within-study randomization. Results are presented in terms of a placebo referent, to rank the drugs based on a common comparator, but data from head-to-head comparisons (e.g., risperidone versus olanzapine) were incorporated in the analysis. An appendix to the report contains the methods and results including those for every possible comparison between the individual drugs. Findings from the network meta-analyses are considered fairly observational in nature and were compared with other more direct findings from the pairwise meta-analyses.

Our primary approach to answer each KQ's parts (a) and (b) on subgroup effects (i.e., variation in effect based on patient and clinical characteristics) was to record any within-study subgroup analyses performed by study investigators using individual patient data; these results preserved 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 also performed our own subgroup analyses using study-level data, where possible. For the benefit outcomes (for which we usually had fewer than 6-10 studies) we performed sensitivity analyses on the results of the pairwise meta-analyses by subgroup variables, such as treatment phase, and/or made observations of the data about possible modification to effect sizes or heterogeneity specific to the subgroup variables of interest. We employed univariate Bayesian meta-regression analyses for four key harm outcomes (weight, weight gain of greater than 7%, somnolence, incidence of any extrapyramidal symptoms) in terms of patient age, sex, antipsychotic treatment history (i.e., % treatment naïve), and treatment duration. We also performed adjusted network meta-analyses using treatment duration (found statistically significant in the metaregression for weight gain) as a study-level variable. These analyses relied on study-level data (e.g., average age in study), such that the results should be considered observational in nature.

Grading the Strength of the Body of Evidence

We followed the Methods Guide and updated guidance[25] to evaluate the strength of the body of evidence for the key outcomes and comparisons. The strength of evidence (SOE) was graded by one reviewer, and reviewed by a second reviewer. Disagreements were resolved through discussion or by consulting with a third reviewer, as needed. Tables of findings were generated for all outcomes and comparisons that had greater than insufficient SOE. We assessed SOE based on five core domains: study limitations, consistency, directness, precision, and reporting bias. For rare events (≤ 5% of patients in both groups having event) we considered 2000 patients sufficient to offer adequate power to detect a difference and therefore provide precise results. For continuous outcomes, more than 400 total enrolled patients are generally considered to offer precise data based on adequate power to detect a 0.2 standardized effect size;[26] we estimated that studies having as few as 200 patients could offer precise estimates of effect. When a confidence interval around an effect estimate was not statistically significant (suggesting no difference) but included values that may be clinically significant for some patients, we could not rule out the possibility of a benefit or harm for this outcome and therefore rated down for precision.

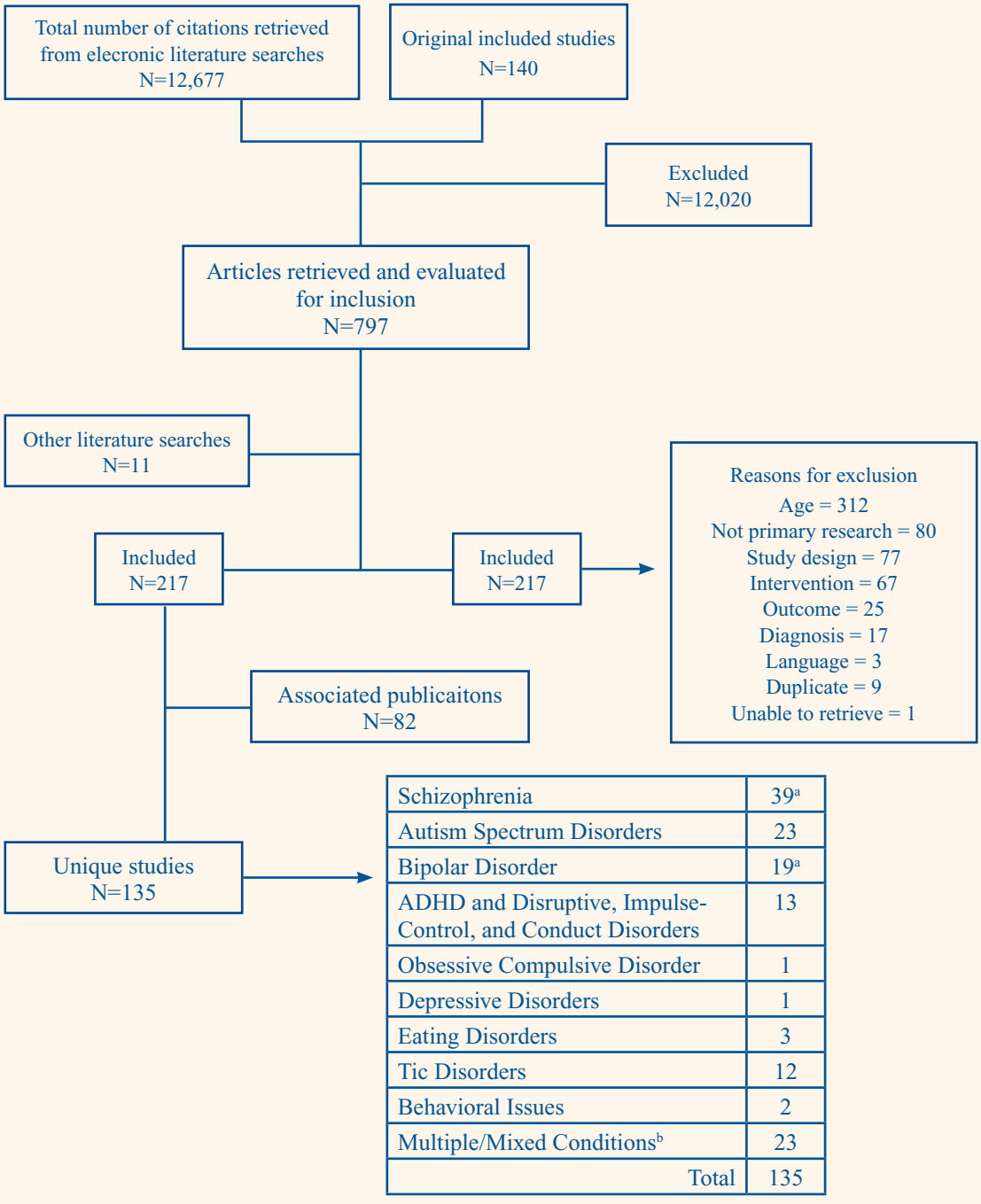
Interpretations of Findings

We chose to use standard wording to describe our interpretations of the SOE and of the magnitude of the effects.[27] For findings supported by high, moderate, low, and insufficient SOE (for which we have similar confidence in the results) we use "will", "probably/likely", "may/appears to", and "not known" in our textual descriptions of the results. Related to magnitude of effects, when the evidence showed effects that would be considered by many patients and practitioners to be clinically important or small, we use "increase/improve/decrease/worsen" (as suitable) or "slightly increase/improve/decrease/worsen", respectively; when there appears to be no difference in effect, we use "makes little or no difference."

Results and Discussion

Our database searches identified 12,677 citations, and 11 additional records were identified from other sources. In total, we included 57 new studies in addition to 78 from the original review (N = 135). Figure B describes the flow of literature through the screening.

Figure B. Flow of literature through study search and selection process



Schizophrenia	39[a]
Autism Spectrum Disorders	23
Bipolar Disorder	19[a]
ADHD and Disruptive, Impulse-Control, and Conduct Disorders	13
Obsessive Compulsive Disorder	1
Depressive Disorders	1
Eating Disorders	3
Tic Disorders	12
Behavioral Issues	2
Multiple/Mixed Conditions[b]	23
Total	135

ADHD = attention deficit hyperactivity disorder

[a] One study provided separate data for both bipolar disorder and schizophrenia; [b]Studies with populations having multiple primary diagnosis were included for key question 2 on harms only.

A total of 100 studies (74%) examined antipsychotics for intermediate and effectiveness outcomes (KQ1). Harms (KQ2) were reported in 126 studies (93%). Of the 135 studies, 89 (66%) were RCTs, 6 (4%) were NRCTs, and 40 (30%) were observational studies.

The number of enrolled/examined participants ranged from 8 to 4140 (median = 59; IQR [interquartile range], 30 to 119). The mean age of study participants ranged from 4 to 22 years (median, 13; IQR, 9.8 to 15.35); studies of schizophrenia generally enrolled older patients (mean 15.8, range 8.86 to 22 years) than those of other conditions (mean 11.34, range 4-19 years). The mean age was lower than 12 years in 52 studies (39%). One hundred and one (75%) studies reported on followup durations of < 6 months, 10 reported on both short- and long-term followup, and 24 reported only on longer-term followup.

Overall, 113 studies provided one or more head-to-head comparisons of individual FGAs or SGAs. A total of 20 studies compared different doses of the same antipsychotic, and 56 studies compared one antipsychotic with placebo. Only five studies included arms with patients taking a variety of SGAs or FGAs.

For subjective outcomes in trials, the overall ROB was rated as high for 60 percent of studies; only eight were assessed as low ROB. The ROB was slightly lower when considering objective outcomes (high for 55% of studies). The main contributor to ROB was incomplete outcome data. Overall, the observational studies were of quite high quality; of 40 studies, 4 (10%) were rated as having high ROB, 12 (30%) as having medium ROB, and 24 (60%) as low ROB. Despite this, the observational studies are still considered of poorer quality (i.e., providing less validity) than the RCTs, because of their inability to completely account for confounding by patient characteristics. Almost half of the studies did not account in some way for variables of confounding considered important (i.e., treatment history, duration/stage of illness).

Key Findings of Intermediate and Effectiveness Outcomes (Key Question 1)

The findings for key intermediate and effectiveness outcomes are summarized below. With the exception of studies examining schizophrenia, the evidence comparing FGAs with SGAs and different antipsychotics within each class was limited. For most conditions, the majority of the findings focused on the comparison of SGA versus placebo. Summary of findings tables contain the findings having at least low SOE.

Schizophrenia and Related Psychoses

Twenty-eight studies reported on intermediate outcomes and 14 reported on effectiveness outcomes for use of FGAs and SGAs in schizophrenia and related psychosis. The average age of patients across the studies was 15.8 years (range 8.9-22). Sexes were fairly equally represented across the studies (60.1% male). Most studies had treatment durations between 4 and 12 weeks; nine studies were 6 months or longer. Table B summarizes the findings.

There may be little or no difference between FGAs and SGAs for the key outcomes of negative symptoms, positive symptoms, response rates, and global impressions of illness severity. The effects for depression symptoms or global impressions of improvement are not known.

Six studies comparing olanzapine with risperidone found that there may be little or no difference in their effects for negative and positive symptoms, response rates, and global impressions of severity. There appears to be little or no difference between low- and high-dose asenapine for response rates or global impressions of severity in the short-term. Between high and low doses of quetiapine, there is probably little or no difference in clinician impressions of severity or global functioning, and there may be little or no difference in reduction in negative symptoms or improvements in response rates. The effects between different doses of other antipsychotics are not known.

Compared with placebo, SGAs as a class likely increase response rates, decrease slightly (not clinically significant) negative and positive symptoms, and improve slightly global impressions of improvement, severity, and functioning. They may make little or no difference in depression symptoms. The only outcome which appeared to result in substantial clinical benefit was response rates (RR, 1.52; 95% CrI, 1.15 to 2.02); the effect estimates for all other outcomes were of a small magnitude, which appears to be influenced by a substantial placebo effect in many cases. Sensitivity analysis by removing the study examining maintenance, rather than acute, treatment with aripiprazole did not affect overall findings to any meaningful extent; results were similar when applying sensitivity analysis for the prodrome phase of psychosis. There appears to be little or no difference between SGAs and placebo for suicide attempts, completed suicide, suicide ideations, or suicide behaviors in short-term studies.

Table B. Summary of findings for schizophrenia and related psychosis: Key intermediate and effectiveness outcomes having at least low strength of evidence

Comparison, Category of Outcome	Outcome (N Studies, N Patients)	Findings,[a] Measurement Tool With Possible Range of Values, if Applicable	Strength of Evidence; Conclusions
SGAs vs. FGAs Intermediate outcomes	Negative symptoms (RCTs: 5, 217)	4 RCTs: SMD, 0.0; 95% CrI, -0.55 to 0.50 1 RCT: No difference (p value NR)	Low; may make little or no difference[b]
	Positive symptoms (RCTs: 5, 217)	4 RCTs: SMD, -0.25; 95% CrI, -0.92 to 0.29 1 RCT: No difference (p value NR)	Low; may make little or no difference[b]
	Response rates (RCTs: 2, 188)	RR, 1.06; 95% CrI, 0.53 to 2.25	Low; may make little or no difference[b]
	Global impressions of severity using CGI-S[c] (RCTs: 2, 124)	MD, -0.21; 95% CrI, -1.19 to 0.67	Low; may make little or no difference[d]
Olanzapine vs. risperidone Intermediate outcomes	Negative symptoms (RCTs: 5, 198)	4 RCTs: SMD, -0.09; 95% CrI, -0.76 to 0.53 1 RCT: No difference p = 0.19	Low; may make little or no difference[b]
	Positive symptoms (RCTs: 5, 198)	4 RCTs: SMD, -0.11; 95% CrI, -0.76 to 0.40 1 RCT: No difference p = 0.10	Low; may make little or no difference[b]
	Response rates (RCTs: 4, 156)	RR, 1.01; 95% CrI, 0.51 to 1.9	Low; may make little or no difference[b]
	Global impressions of severity using CGI-S (RCTs: 3, 131)	1 RCT: MD, 0.30; 95% CI, -0.53 to 1.13 1 RCT: MD, 0.30; 95% CI, -0.41 to 1.01 1 RCT: No difference p = 0.33	Low; may make little or no difference[d]
Asenapine high vs. low	Response rate (RCTs: 1, 204)	1 RCT: RR, 1.00; 95% CI, 0.75 to 1.32	Low; may make little or no difference[e]
	Global impressions of severity using CGI-S (RCTs: 1, 204)	1 RCT: MD, 0.20; 95% CI, -0.05 to 0.45	Low; may make little or no difference[e]
Quetiapine high vs. low dose Intermediate outcomes	Negative symptoms (RCTs: 2, 238)	1 RCT: MD, 1.6; 95% CI, -4.79 to 7.99 (SANS; range 0-25) 1 RCT: MD, 0.14; 95% CI, -1.81 to 2.09 (PANSS; range 7-49)	Low; may make little or no difference[b]
	Response rates (RCTs: 2, 273)	1 RCT: RR, 0.73; 95% CI, 0.41 to 1.29 1 RCT: RR, 1.05; 95% CI, 0.69 to 1.60	Low; may make little or no difference[b]
	Global impressions of severity using CGI-S (RCTs: 2, 238)	1 RCT: MD, 0.00; 95% CI, -0.35 to 0.35 1 RCT: MD, -0.13; 95% CI, -0.47 to 0.21	Moderate; probably makes little or no difference[f]
	Global impressions of functioning (RCTs: 2, 238)	1 RCT: MD, -3.5; 95% CI, -8.37 to 1.37 (GAF; range 1-100) 1 RCT: MD, 1.9; 95% CI, -2.35 to 6.15 (C-GAS; range 1-100)	Moderate; probably makes little or no difference[f]

Table B. Summary of findings for schizophrenia and related psychosis: Key intermediate and effectiveness outcomes having at least low strength of evidence (continued)

Comparison, Category of Outcome	Outcome (N Studies, N Patients)	Findings,[a] Measurement Tool With Possible Range of Values, if Applicable	Strength of Evidence; Conclusions
All SGAs vs. placebo Intermediate outcomes	Negative symptoms (RCTs: 9, 1788)	MD, -1.31; 95% CrI, -2.05 to -0.58 (PANSS Negative; range 7-49)	Moderate; SGAs probably decrease slightly[f]
	Positive symptoms (RCTs: 9, 1788)	MD, -2.20; 95% CrI, -2.98 to -1.48 (PANSS Positive; range 7-49)	Moderate; SGAs probably decrease slightly[f]
	Depression symptoms (RCTs: 2, 420)	1 RCT: MD, -0.59; 95% CI, -1.46 to 0.28 1 RCT: MD, -0.59; 95% CI, -1.45 to 0.27 (PANSS Depression)	Low; may make little or no difference[f]
	Response rates (RCTs: 5, 993)	RR, 1.52; 95% CrI, 1.15 to 2.02	Moderate; SGAs probably increase[f]
	Global impressions of improvement using CGI-I (RCTs: 6, 1202)	MD, 0.54; 95% CrI, -1.07 to -0.14	Moderate; SGAs probably improve slightly[f]
	Global impressions of severity using CGI-S (RCTs: 9, 1788)	MD, -0.36; 95% CrI, -0.51 to -0.22	Moderate; SGAs probably improve slightly[f]
	Global impressions of functioning (RCTs: 7, 1339)	MD, 4.15; 95% CrI, 2.03 to 6.59 (C-GAS; range 0-100)	Moderate; SGAs probably improve slightly[f]
All SGAs vs. placebo Effectiveness Outcomes	Short-term suicide attempts/ suicides (RCTs: 7, 1463)	Attempts: 2 in 693 SGA and 2 in 318 placebo patients Suicides: 0 in 447 SGA vs. 0 in 227 placebo patients	Low; may make little or no difference[g]
	Short-term suicide ideations or behaviors (RCTs: 4, 758)	Ideations: 3 in 340 SGA and 1 in 165 placebo patients Behaviors: 1 in 170 SGA and 1 in 83 placebo patients	Low; may make little or no difference[g]

C-GAS = Global Assessment Scale for Children; CGI-I = Clinical Global Impressions of Improvement; CGI-S = Clinical Global Impressions of Severity; CI = confidence interval; CrI = credible interval (used with Bayesian meta-analysis); FGA = first-generation antipsychotic; GAF = Global Assessment of Functioning; MD = mean difference; N = number; NR = not reported; PANSS; Positive and Negative Syndrome Scale; RCT: randomized controlled trial; ROB = risk of bias; RR = risk ratio; SANS = Scale for the Assessment of Negative Symptoms; SGA = second-generation antipsychotics; SMD = standardized mean difference

[a] When the findings are not representative of the total number of studies identified in the outcome column, we included the number of studies for each finding; we did not pool data from 1 or 2 studies so these results are always presented separately. All values except Response and Global Impressions of Functioning are favorable for group 1 (G1) when there is a negative effect estimate; the larger the magnitude of the number the larger the effect. SMDs provide results in standard deviation units, and are used when the results from different measurement tools are combined in meta-analysis; as a general rule, 0.2 represents a small effect size, 0.5 a moderate one, and 0.8 a large one.

[b] Downgraded for ROB and imprecision, because credible interval includes a clinically significant value (e.g., SMD ≥ ±0.50, CGI-I or CGI-S ≥ ±2 points [7 point scales]) such that we could not rule out benefit even though effect estimate appears to be of no difference.

[c] CGI-S and CGI-I scores range from 0-6.

[d] Downgraded for ROB and imprecision because of small sample size, typically < 200 patients in total.

[e] Downgraded for inconsitnecy and imprecision.

[f] Downgraded for ROB.

[g] Downgraded for ROB and imprecision because of small event rates; confidence intervals of relative risks ranged between 0.02 to 5.0, to 0.06 to 48.1).

Bipolar Disorder

Of 19 studies examining treatment of bipolar disorder, 15 reported on intermediate and 11 on effectiveness outcomes. The average age of patients was 12.8 years. Both sexes were equally represented across the studies (56% male). Sixteen trials had followup periods ranging from 3 to 12 weeks. One trial had a controlled extension phase of 30 weeks, one trial had a placebo-controlled maintenance treatment duration of 72 weeks, and an observational study reviewed charts for between 7 to 8 months. Table C contains a summary of the findings.

There may be a slightly greater reduction in manic symptoms from high- (10mg/day) versus low-dose (5 mg/day) asenapine; dose of asenapine may make little or no difference for global impressions of severity or for depression.

Compared with placebo, SGAs likely reduce manic symptoms and probably decrease slightly depression symptoms. SGAs probably increase response and remission rates versus placebo in studies of patients experiencing manic/mixed phases; clinical and statistical heterogeneity was introduced when including two RCTs examining quetiapine for patients with depressive episodes (showing less response). Moderate SOE exists showing that SGAs probably decrease symptom severity to a small extent and increase global functioning slightly compared with placebo.

When examining individual SGAs versus placebo, the findings for aripiprazole were similar to those across all SGAs, with the exception of depression symptoms where use of this SGA may make little or no difference. Quetiapine probably reduces manic symptoms, likely makes little or no difference for depression symptoms, and appears to make no difference for response in studies of patients experiencing manic/mixed episodes; the results of little to no difference for response rates (often focused on manic symptoms) were imprecise showing that many patients may have clinically relevant response. The effects of quetiapine versus placebo for remission rates and for global impressions of severity are not known.

A study enrolling patients with prodromal bipolar disorder reported similar efficacy to the other studies of patients with manic symptoms. A study exclusively enrolling patients having comorbid ADHD did not appear to differ in effect for several outcomes to other similar studies assessing SGAs in manic or mixed episodes. Several within-study subgroup analyses showed that concomitant use of psychostimulants had no significant effect on manic symptoms; comorbid diagnosis of ADHD or a disruptive, impulse-control, or conduct disorder did not significantly affect results either for mania or depression.

For effectiveness outcomes, SGAs may make little or no difference over placebo for suicide ideations and attempts.

Table C. Summary of findings for bipolar disorder: Key intermediate and effectiveness outcomes having at least low strength of evidence

Comparison, Outcome Category	Outcome (N Studies; N Patients)	Findings,[a] Tool With Range of Values, if Applicable	Strength of Evidence; Conclusions
Asenapine high (10 mg/day) vs. low (5 mg/day) dose	Manic symptoms (1, 199)	MD, -2.80; 95% CI -0.64 to -4.96 (YMRS; range 0-60)	Low; High-dose asenapine may decrease slightly manic symptoms[b]
	Global impressions of severity (1, 199)	MD, -0.10, 95% CI -0.29 to 0.49	Low; may make little or no difference[b]
	Depression (1, 199)	MD, 0.80; 95% CI -1.87 to 3.47 (CDRS; range 0-113)	Low; may make little or no difference[b]
All SGAs vs. placebo Intermediate Outcomes	Manic symptoms (11, 1639)	MD, -6.42; 95% CrI, -7.88 to -5.26 (YMRS; range 0-60)	Moderate; SGAs probably decrease[c]
	Depression symptoms (9, 1622)	MD, -1.65; 95% CrI, -2.78 to -0.48 (CDRS; range 0-113)	Moderate; SGAs probably decrease slightly[c]
All SGAs vs. placebo Intermediate Outcomes (continued)	Response (10, 1664) (Manic/mixed phases)[d]	RR, 1.97; 95% CrI, 1.66 to 2.34 (40-50% reduction in YMRS from baseline)	Moderate; SGAs probably increase for manic/mixed phases[c]
	Remission (5, 944) (Manic/Mixed phases)[d]	RR, 2.84; 95% CrI, 1.67 to 5.55	Moderate; SGAs probably increase for manic/mixed phases[c]
	Global impressions of severity using CGI-S[e] (9, 1778)	MD, -0.65; 95% CI, -0.80 to -0.49	Moderate; SGAs probably slightly decrease[c]
	Global impressions of functioning (4, 1188)	MD, 6.64; 95% CrI, 2.45 to 10.95 (C-GAS; range 1-100)	Moderate; SGAs probably slightly increase[c]
All SGAs vs. placebo Effectiveness Outcomes	Suicide ideation (8, 1782)	RR, 1.12; 95% CrI, 0.58 to 2.26	Low; SGAs may make little or no difference[f]
	Suicide attempts (6, 1285)	RR, 1.71; 95% CrI, 0.39 to 7.38	Low; SGAs may make little or no difference[f]
Aripiprazole vs. placebo Intermediate Outcomes	Manic symptoms (3, 387)	MD, -7.08; 95% CrI, -10.96 to -3.24 (YMRS; range 0-60)	Moderate; Aripiprazole probably decreases[c]
	Depression symptoms (2, 311)	1 RCT: MD, -1.74; 95% CI, -3.92 to 0.44 1 RCT: MD, -2.29; 95% CI, -10.62 to 6.04 (CDRS-R; range 17-113)	Low; Aripiprazole may make little or no difference[g]
	Response rates (2, 311)	1 RCT: RR, 2.11; 95% CI, 1.47 to 3.02 1 RCT: RR, 1.71; 95% CI, 1.13 to 2.58	Moderate; Aripiprazole probably increases[c]

Table C. Summary of findings for bipolar disorder: Key intermediate and effectiveness outcomes having at least low strength of evidence (continued)

Comparison, Outcome Category	Outcome (N Studies; N Patients)	Findings,[a] Tool With Range of Values, if Applicable	Strength of Evidence; Conclusions
Aripiprazole vs. placebo Intermediate Outcomes (continued)	Remission (2, 311)	1 RCT: RR, 7.09; 95% CI, 2.96 to 16.99 1 RCT: RR, 2.26; 95% CI, 1.19 to 4.28	Moderate; Aripiprazole probably increases[c]
	Global impressions of severity using CGI-S (2, 328)e	1 RCT: MD, -1.00; 95% CI, -1.34 to -0.67 1 RCT: MD, -0.41; 95% CI, -0.80 to -0.02	Moderate; Aripiprazole probably slightly decreases[c]
Quetiapine vs. placebo Intermediate Outcomes	Manic symptoms (3, 339)	MD, -5.34; 95% CrI, -9.92 to -0.44 (YMRS; range 0-60)	Moderate; Quetiapine probably decreases[c]
	Depression symptoms (3, 501)	MD, -1.87; 95% CrI, -4.71 to 1.11 (CDRS-R; range 17-113)	Moderate; Quetiapine probably makes little or no difference[c]
	Response (2, 307) (Manic/mixed)	1 RCT: RR, 1.36; 95% CI, 0.97 to 2.72 1 RCT: RR, 1.97; 95% CI, 1.38 to 2.81	Low; Quetiaipine may make little or no difference[g]

CDRS-R = Children's Depression Rating Scale-Revised; C-GAS = Global Assessment Scale for Children; CGI-S = Clinical Global Impressions of Severity; CrI = credible interval (used with Bayesian meta-analysis); MD = mean difference; N = number; RCT: randomized controlled trial; RR = risk ratio; SGA = second-generation antipsychotics; YMRS = Young Mania Rating Scale

[a] When the findings are not representative of the total number of studies identified in the outcome column, we included the number of studies for each finding; we did not pool data from 1 or 2 studies so these results are always presented separately. All values except Response, Remission, and Global Impressions of Functioning are favorable for the SGA when there is a negative effect estimate; the larger the magnitude of the number the larger the effect.

[b] Dowgraded for imprecision.

[c] Downgraded for ROB.

[d] When two studies examining the depressive phase were included the heterogeneity has substantial.

[e] CGI-S scores range from 0-6.

[f] Downgraded for ROB and imprecision due to small samples for this rare outcome.

[g] Downgraded for ROB and imprecision due to CI including clinically relevant benefit for SGAs.

Autism Spectrum Disorders

Twenty-three studies examined the effectiveness of FGAs and SGAs in autism spectrum disorders. The average age of patients was 9.1 years, and patients were predominantly male (average 83%). Treatment duration varied widely across studies (range, 4 weeks to 2.3 years). For the studies (n = 18) we considered short-term (< 6 months duration), average duration was 8.9 weeks. Table D summarizes the findings.

At least low SOE was only found for intermediate outcomes in comparisons between SGA and placebo. SGAs probably decrease irritability, and probably slightly decrease lethargy/social withdrawal, stereotypy, and inappropriate speech. SGAs likely increase response rates and (slightly) clinical severity. They may increase global impressions of improvement. Maintenance treatment with an SGA appears to decrease relapse rates.

When examining studies of aripiprazole and risperidone, the findings were similar for irritability and (with aripiprazole) for stereotypy. For lethargy, inappropriate speech, and response rates (with risperidone) conclusions were that these SGAs may make little or no difference; smaller sample sizes contributing to the SOE for each drug likely affected the ability to obtain a significant finding for most outcomes (e.g., response rates), with the exception of irritability which overall had the larger magnitude of effect.

Table D. Summary of findings for autism spectrum disorders: Key intermediate outcomes having at least low strength of evidence

Comparison	Outcome (N Studies; N Patients)	Findings,[a] Tool With Range of Values, if Applicable	Strength of Evidence; Conclusion
SGAs vs. placebo	Irritability (8, 809)	MD, -6.38; 95% CrI, -8.94 to -3.83 (ABC subscale; range 0-45)	Moderate; SGAs probably decrease[b]
	Lethargy/social withdrawal (7, 743)	MD, -1.67; 95% CrI, -3.05 to -0.28 (ABC subscale; range 0-48)	Moderate; SGAs probably decrease slightly[b]
	Stereotypy (5, 634)	MD, -1.73; 95% CrI, -3.16 to -0.05 (ABC subscale; range 0-21)	Moderate; SGAs probably decrease slightly[b]
	Inappropriate speech (7, 743)	MD, -1.04; 95% CrI, -1.83 to -0.26 (ABC subscale; range 0-12)	Moderate; SGAs probably decrease slightly[b]
	Response rates (7, 716)	RR, 2.22; 95% CrI, 1.29 to 4.17	Moderate; SGAs probably increase[b]
	Relapse rates (3, 141) (Maintenance phase only)	RR, 0.30; 95% CrI, 0.07 to 0.84	Low; SGAs may decrease during maintenance treatment[c]
	Global impressions of improvement on CGI-I[d] (6, 635)	4 RCTs: MD, -1.00, 95% CrI, -2.34 to 0.07 3 RCTs: RR 4.5 and 6.5; both $p < 0.01$ (proportion scoring as at least "much improved")	Low; SGAs may increase[b]
	Global impressions of severity on CGI-S[d] (4, 522)	3 RCTs: MD, -0.61; 95% CrI, -1.04 to -0.15	Moderate; SGAs probably slightly decrease[b]
Aripiprazole vs. placebo	Irritability (3, 393)	MD, -5.74; 95% CrI, -9.34 to -2.15 (ABC subscale; range 0-45)	Moderate; Aripiprazole probably decreases[b]
	Lethargy/social withdrawal (3, 393)	MD, -1.41; 95% CrI, -4.19 to 1.35 (ABC subscale; range 0-48)	Low; Aripiprazole may make little or no difference[e]
	Stereotypy (3, 393)	MD, -2.51; 95% CrI, -4.68 to -0.33 (ABC subscale; range 0-21)	Moderate; Aripiprazole probably decreases slightly[b]
	Inappropriate speech (3, 393)	MD, -1.49; 95% CrI, -3.02 to 0.06 (ABC subscale; range 0-12)	Low; Aripiprazole may make little or no difference[e]
Risperidone vs. placebo	Irritability (4, 268)	MD, -8.28; 95% CrI, -12.59 to -3.64 (ABC subscale; range 0-45)	Moderate; Risperidone probably decreases[b]
	Lethargy/social withdrawal (3, 202)	MD, -2.51; 95% CrI, -5.67 to 1.02 (ABC subscale; range 0-48)	Low; Risperidone may make little or no difference[e]
	Stereotypy (2, 178) (Acute phase only)	1 RCT: -3.10; 95% CI, -4.93 to -1.27 1 RCT: -1.90; 95% CI, -3.64 to -0.16 (ABC subscale; range 0-21)	Low; Risperidone may decrease slightly in acute treatment[c]
	Inappropriate speech (3, 202)	MD, -1.06; 95% CrI, -2.66 to 0.59 (ABC subscale; range 0-12)	Low; Risperidone may make little or no difference[e]
	Response rate (3, 246)	RR, 2.75; 95% CrI, 0.92 to 9.77	Low; Risperidone may make little or no difference[e]

ABC = Aberrant Behavior Checklist; CB-YOCS = Children's Yale-Brown Obsessive Compulsive Scale; CGI-I = Clinical Global Impressions of Improvement; CGI-S = Clinical Global Impressions of Severity; CI = confidence interval; CrI = credible interval (used with Bayesian meta-analysis); MD = mean difference; N = number; RCT: randomized controlled trial; RR = risk ratio; SGA = second-generation antipsychotics

[a] When the findings are not representative of the total number of studies identified in the outcome column, we included the number of studies for each finding; we did not pool data from 1 or 2 studies so these results are always presented separately. All values except Response are favorable for SGAs when there is a negative MD, or a RR < 1.0 (i.e., relapse); the larger the magnitude of effect, the larger the effect.

[b] Downgraded for ROB.

[c] Downgraded for ROB and imprecision because of small sample size, typically < 200 patients in total.

[d] CGI-S and CGI-I scores range from 0-6.

[e] Downgraded for ROB and imprecision, because credible interval includes a clinically significant value (e.g., lower boundary value considered clinically meaningful reduction) such that we could not rule out benefit even though effect estimate appears to be of no difference.

ADHD and Disruptive, Impulse-Control, or Conduct Disorders

Thirteen studies examined ADHD and/or disruptive, impulse-control, or conduct disorders (DICD). Patients had an average age of 9.9 years and were predominantly male (83%); apart from two RCTs enrolling adolescents, the age of participants was typically below 12 years and close to 9-10 years (no study had a mean age below 8 years). Most RCTs were examining acute phase treatment in patients either naïve to or not taking antipsychotics upon enrollment; one RCT enrolled children maintained on risperidone for 1 year and examined placebo-controlled discontinuation of the antipsychotic. All children were taking stimulants in three RCTs, variable numbers were taking stimulants in five RCTS, and stimulants were prohibited in three RCTs. We summarize the findings in Table E. All evidence graded as having at least low SOE was for outcomes between SGAs and placebo.

Compared with placebo, SGAs as a class (and risperidone alone) probably reduce conduct problems and aggression in children with ADHD and/or DICD. Results for clinical impressions of improvement showed little or no difference, although results were imprecise and indicated that many patients may possibly improve. Risperidone likely decreases hyperactivity, although this level of confidence is specific to studies where not all patients are taking, or are not responding to, stimulant medications. SGAs (and risperidone) appear to reduce clinical severity, and they probably reduce severity more for patients with a primary diagnosis of DICD rather than ADHD. Studies found that SGAs may make little or no difference compared with placebo for global impression of improvement. From two RCTs of patients with primarily ADHD and aggression, risperidone appears to make little or no difference for response rates.

From between-study observations, risperidone may preferentially reduce illness severity, and increase global improvement ratings, for primary diagnosis of DICD compared with ADHD particularly when used for ADHD as adjunctive treatment. Our meta-analysis favored risperidone over placebo for hyperactivity, although the data came from studies where not all patients were taking stimulants, or to the situation of nonresponse to stimulants; a study with children responding to stimulants found no benefit for risperidone on hyperactivity. Sensitivity analyses for the small study enrolling children with a history of response to risperidone did not affect the results. We did not find any evidence to suggest a differential treatment effect based on patients' intellectual functioning.

Five studies of ADHD and DICD conducted analyses of outcomes in different subpopulations. Two studies found no effect of age for effects of risperidone on aggression or risk of symptom recurrence. One RCT found no impact of comorbidities (including global developmental delay, ADHD, and secondary diagnosis of disruptive behavior disorders) or cotreatment with psychostimulants on conduct problems. A pooled analysis of two similar RCTs found no indication that the effects of risperidone on conduct problems or hyperactivity varied with stimulant use. Risperidone-naïve patients had lower conduct problem scores in one study, whereas prior treatment had no impact on symptom severity in another study.

Table E. Summary of findings for ADHD and disruptive, impulse-control, or conduct disorders: Key intermediate outcomes having at least low strength of evidence

Comparison	Outcome (N Studies; N Patients)	Findings[a]	Strength of Evidence; Conclusion
SGAs vs. placebo	Conduct problems (6, 462)	SMD, -0.77; 95% CrI, -1.34 to -0.17	Moderate; SGAs probably decrease[b]
	Aggression (7, 495)	SMD, -0.43; 95% CrI, -0.67 to -0.14	Moderate; SGAs probably decrease[b]
	Global impressions of improvement using CGI-I[c] (7, 482)	5 RCTs: RR, 2.13; 95% CrI, 0.87 to 6.46 (proportion at least "improved") 1 RCT: MD, -0.50; 95% CI, -1.99 to 0.99 1 RCT: MD, -1.80; 95% CI, -2.89 to -0.71	Low; SGAs may make little or no difference[d]
	Global impressions of severity using CGI-S (3, 75) (Studies of primary treatment in DICD)	MD, -1.98; 95% CrI, -3.18 to -0.93	Low; SGAs may reduce in DICD[e]
Risperidone vs. placebo	Conduct problems (5,443)	SMD, -0.84; 95% CrI, -1.54 to -0.18	Moderate; Risperidone probably decreases[b]
	Aggression (6, 476)	SMD, -0.44; 95% CrI, -0.72 to -0.13	Moderate; Risperidone probably decreases[b]
	Hyperactivity (6, 468) (Specific to primary diagnosis of DICD and study of those with ADHD not responding to stimulants)	5 RCTs: SMD, -0.39; 95% CrI, -0.76 to -0.07 1 RCT: No difference $p > 0.05$ (All patients taking stimulants)	Moderate; Risperidone probably decreases for those with primary diagnosis of DICD or ADHD if not responding to stimulants[b]
	Global impressions of improvement using CGI-I (6, 463)	4 RCTs: RR, 1.85; 95% CrI, 0.64 to 5.58 (proportion at least "improved") 1 RCT: MD, -0.50; 95% CI, -1.99 to 0.99 1 RCT: MD, -1.80; 95% CI, -2.89 to -0.71	Low; Risperidone may make little or no difference[d]
	Global impressions of severity using CGI-S (2, 56) (Studies of primary treatment in DICD)	1 RCT: MD, -1.80; 95% CI, -2.54 to -1.06 1 RCT: MD, -2.50; 95% CI, -4.11 to -0.89	Low; Risperidone may decrease in DICD[e]
	Global impressions of severity using CGI-S (2, 193) (Studies of stimulant augmentation in ADHD)	1 RCT: MD, 0.0; 95% CI, -1.65 to 1.65 1 RCT: RR, 1.2; 95% CI, 0.95 to 1.5 (proportion rated as "normal/borderline/mildly ill")	Low; Risperidone may make little or no difference in ADHD treatment augmented with risperidone[d]
	Response rate (2, 193) (Patients with primarily ADHD and aggression)	1 RCT: RR, 1.12; 95% CI, 0.94 to 1.34 1 RCT: RR, 1.28; 95% CI, 0.93 to 1.77	Low; Risperidone may make little or no difference in patients with primary diagnosis of ADHD and aggression[d]

ADHD = attention-deficit/hyperactivity disorder; CGI-S = Clinical Global Impressions of Severity; CI = confidence interval; CrI = credible interval (used with Bayesian meta-analysis); DICD = disruptive, impulse-control, and conduct disorders; MD = mean difference; N = number; RCT: randomized controlled trial; RR = risk ratio; SGA = second-generation antipsychotics

[a] When the findings are not representative of the total number of studies identified in the outcome column, we included the number of studies for each finding; we did not pool data from 1 or 2 studies so these results are always presented separately. All effect estimates reported as MD or SMD values favor SGAs when they are negative (larger magnitude greater effect); a RR >1.0 favor SGAs. SMDs provide results in standard deviation units, and are used when the results from different measurement tools are combined in meta-analysis; as a general rule, an absolute magnitude of 0.2 represents a small effect size, 0.5 a moderate one, and 0.8 a large one.

[b] Downgraded for ROB.

[c] CGI-S and CGI-I scores range from 0-6.

[d] Downgraded for ROB and imprecision, because credible interval includes a clinically significant value (e.g., RR ≤0.75 or ≥1.25) such that we could not rule out benefit even though effect estimate appears to be of no difference.

[e] Downgraded for ROB and impression due to small sample size

Obsessive-Compulsive Disorder

One 12-week RCT with 79 patients examined augmentation with risperidone or aripiprazole in patients with obsessive-compulsive disorder (OCD) who failed to respond to at least 12 weeks of treatment with selective serotonin reuptake inhibitors. No significant differences were found between risperidone and aripiprazole for nonspecific symptoms (i.e., response rates were 51.4% and 61.8% for risperidone and aripiprazole, respectively), and global impressions of severity and functioning. Results for core symptoms of obsessions and compulsions were not reported by the authors. All patients had comorbid tic disorders; response to tic symptomatology was similar with 68 percent in both groups responding. Because of insufficient SOE, the effects of risperidone or aripiprazole augmentation of SSRIs in OCD is not known.

Depression

One observational study examined a subgroup of 35 patients aged ≤ 25 years in a pooled analysis of data from two RCTs of placebo-controlled adjuvant aripiprazole (2-20 mg/day) for patients with major depressive disorder who failed to respond to 8 weeks of antidepressant treatment. The focus of the report was on suicidality. Findings suggested no differences in suicidality between placebo and aripiprazole for adjuvant treatment of SSRIs, but we have no confidence in these findings (insufficient SOE).

Eating Disorders

Two RCTs and one retrospective cohort study examined SGAs versus placebo for adjunctive treatment in eating disorders. All three studies enrolled females (average ages 14-18) with anorexia nervosa or eating disorders not-otherwise specified (allowing for persistence of menstruation), who were also receiving multidisciplinary, tailored care within eating disorder programs. Trials of olanzapine and risperidone compared with placebo failed to demonstrate any benefit from these SGAs in terms of increased body weight (favorable for this condition) or reduced eating disorder symptomatology. Findings from the observational study were substantially confounded by a greater illness severity and overall resource use by the olanzapine group. Speculated changes in resting energy expenditure were not realized. The SOE was graded as insufficient for all key outcomes (i.e., weight) of relevance. The studies did not report any effectiveness outcomes.

Tic Disorders

Twelve trials studies tic disorders. All but one study enrolled patients with Tourette's syndrome. Patients enrolled in the studies had an average age of 10.7 years and were predominantly male (84%). Patients had a variety of comorbidities, including ADHD (34%); obsessive-compulsive disorder (23%); and disruptive, impulse-control, and conduct disorders (5%). Only one study permitted concomitant psychotropic medications including stimulants. Table F summarizes the findings for outcomes having at least low SOE.

Tic severity may be reduced in patients receiving SGAs (aripiprazole, risperidone, and ziprasidone). A 6-point reduction in tic severity using the Yale Global Tic Severity Scale's total tic score has empirical evidence of clinical significance.[28]

Table F. Summary of findings for tic disorders: Key intermediate outcomes having at least low strength of evidence

Comparison	Outcome (N studies; N patients)	Findings,[a] Tool With Range of Values	Strength of Evidence; Conclusion
SGAs vs. placebo	Tic severity (3, 114)	MD, -6.26; 95% CrI, -10.05 to -2.54 YGTSS Total Tic score (range 0-50)	Low; SGAs may decrease[b]

CrI = credible interval; N = number; MD = mean difference; ROB = risk of bias; SGA = second-generation antipsychotic; YGTSS = Yale Global Tic Severity Scale.

[a] A negative MD score favors the SGAs.

[b] Downgraded for ROB and imprecision because of small sample size (typically < 200 patients).

Behavioral Issues

Two 4-week RCTs compared risperidone with placebo for treatment of behavioral issues in children without psychiatric diagnoses within this review's condition categories. The inclusion criteria in one study (N = 13) were persistent behavioral disturbances (e.g., hostility, aggressiveness, irritability, agitation) in children with intellectual impairment living in residential homes. Compared with placebo, risperidone significantly reduced symptoms of irritability and hyperactivity, but not lethargy, stereotypic behavior, or inappropriate speech; ratings of clinical improvement were also superior for risperidone.

The other study (N = 90) focused on children diagnosed clinically as having a masturbation problem. Risperidone reduced the frequency of masturbation compared with no medication.

All key outcomes were assessed as having insufficient SOE, therefore the effects in all cases are not known.

Key Findings for Harms Across Conditions (Key Question 2)

This section presents the evidence from analyses across all comparisons for the outcomes of weight and BMI, and then for all key outcomes for head-to-head and then placebo-controlled comparisons. Within each comparison, we begin with findings for major adverse effects (AEs) followed by general AEs. Limited evidence was provided for FGAs. The majority of the findings focused on the comparison of SGA versus placebo. The section ends with findings from subgroup analyses.

All Comparisons: Network Meta-Analyses for Body Composition Outcomes

We conducted network meta-analyses for the outcomes of weight and BMI. These outcomes represent two of the key outcomes that were reported by the most studies (weight, n = 71; BMI, n = 35). We used data regardless of followup duration and (for those with multiple timepoints) from each study's longest term followup; 14 studies for weight and 11 for BMI reported data for treatment durations 6 months or longer. Findings from our analyses are presented in Figures C and D. Results are presented in terms of a placebo referent, to rank the drugs based on a common comparator, but data from head-to-head comparisons were incorporated in the analysis. An appendix to the full report contains the results for every possible comparison between the individual drugs.

Results showed that patients taking most antipsychotics gain more weight than patients taking placebo or not receiving antipsychotics. Molindone and ziprasidone may cause less weight gain on average whereas those receiving olanzapine may gain as much as 5 kilograms more weight during treatment durations of a relatively short timeframe (81% of studies for this analysis were short-term which was often 6-12 weeks duration). Not all SGAs appear to contribute to more weight gain than FGAs. Results for olanzapine clearly separated this SGA as more harmful than most other SGAs. Some of the antipsychotics (e.g., pimozide, molindone, lurasidone) had few patients contributing to the findings which resulted in wide credible intervals. The relative harm from olanzapine is most robust compared with aripiprazole, quetiapine, and risperidone because of the precision in these estimates from larger sample sizes.

For BMI, olanzapine and clozapine were worst for average effect, although the results for clozapine and lurasidone are considerably imprecise because of small samples. Seventy-one percent of studies had short-term treatment durations.

Figure C. Plot of network meta-analysis results for weight gain compared with reference standard (placebo/no treatment)

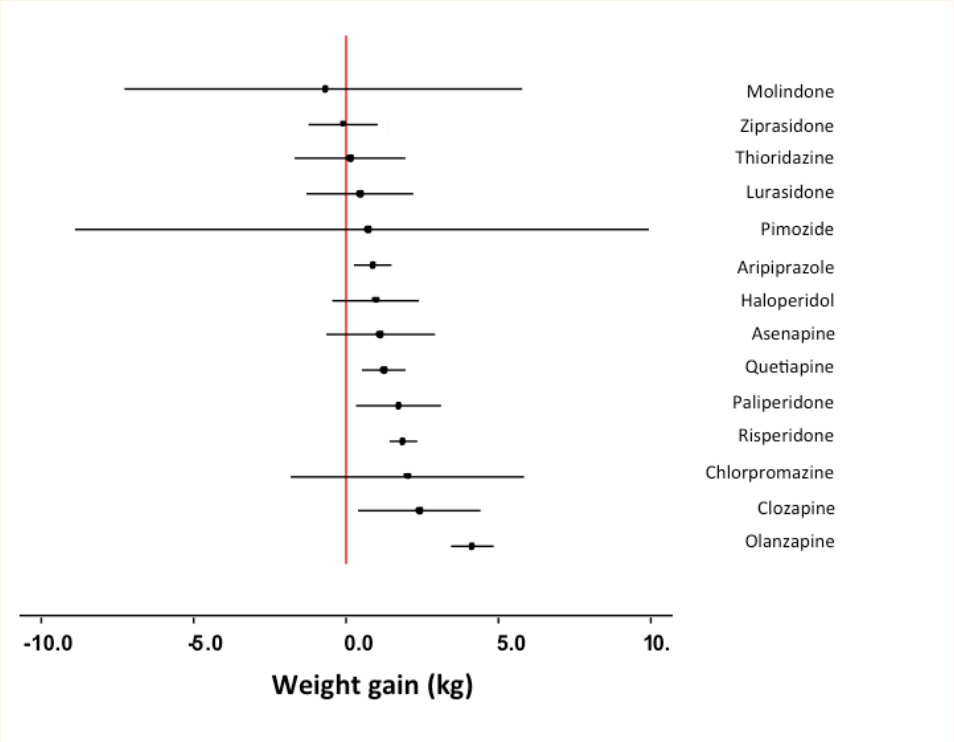
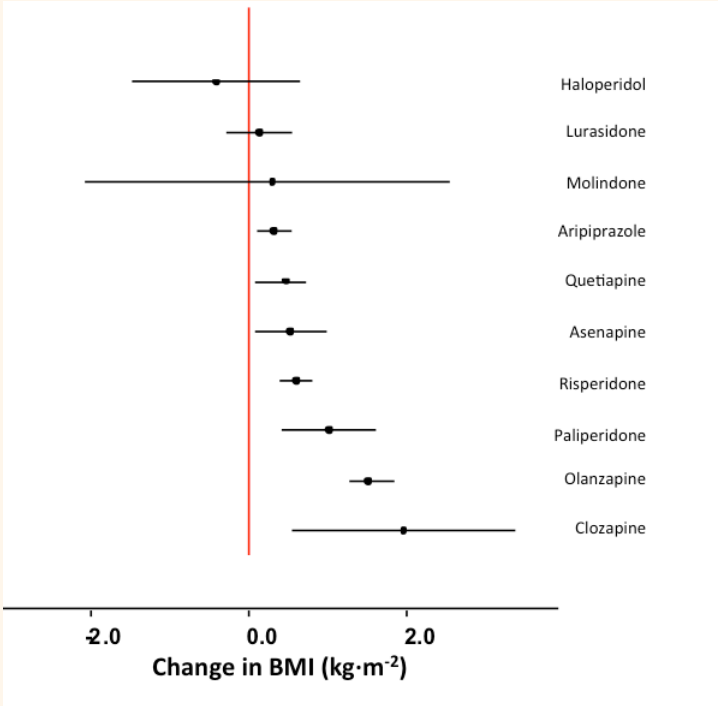


Figure D. Plot of network meta-analysis results for increase in body mass index (BMI) compared with reference standard (placebo/no treatment)



These plots show the findings from network meta-analyses combining placebo-controlled and head-to-head comparisons of first-generation antipsychotics and second-generation antipsychotics within one analysis. The effects shown represent the mean difference and credible intervals of each drug relative to placebo which was used as the reference standard.

FGAs Versus SGAs

Nine studies reported on major (4 long-term duration) and 16 reported on general AEs (2 long-term). Few studies having small sample sizes reported on major AEs which were often rare outcomes. The difference in effects between SGAs and FGAS for all major AEs are not known (insufficient SOE). Table G contains a summary of our key findings for general AEs which are limited to findings of short treatment durations.

Compared with FGAs, SGAs may decrease the risk for experiencing any extrapyramidal symptom (EPS). FGAs probably cause lower gains in weight and BMI. There may be little or no difference between classes for sedation. Evidence was insufficient for other outcomes (e.g., akathisia, dystonia, hyperprolactinemia).

Table G. Summary of findings for general adverse effects: Short-term durations of FGAs versus SGAs

Outcome	N Studies, N Patients	FGA Events	FGA N	SGA Events	SGA N	Relative Effects[a]	Strength of Evidence; Conclusion
Any EPS	4, 110	16	37	13	73	RR, 2.59; 95% CrI, 1.00 to 7.00	Low; SGAs may decrease risk[b]
Weight (kg)	14, 506	-	190	-	316	MD, -2.62; 95% CrI, -4.35 to -0.86	Moderate; FGAs probably better[c]
BMI ($kg.m^{-2}$)	7, 236	-	73	-	163	MD, -1.57; 95% CrI, -2.49 to -0.53	Moderate; FGAs probably better[c]
Sedation	7, 345	70	160	79	185	RR, 1.04; 95% CrI, 0.86 to 1.37	Low; may be little or no difference[d]

AE = adverse effect; BMI = body mass index; CrI = credible interval; FGA = first-generation antipsychotic; G = group; kg = kilogram; m = meter; MD = mean difference; N = number; RR = risk ratio; SGA = second-generation antipsychotic

aRisk ratios above 1.0 and positive MD favor SGAs.

bDowngraded for ROB and imprecision, based on small sample size.

cDowngraded for ROB.

dDowngraded for ROB and imprecision, because CrI includes possibility for clinically relevant benefit for SGAs.

FGAs Versus FGAs

Two short-term RCTs reported on major AEs and provided insufficient SOE for all outcomes. No findings for general AEs in comparisons of FGAs versus FGAs, or between different doses of FGAs, were rated as at least low SOE.

SGAs Versus SGAs: Comparison of Different Drugs

Sixteen (5 long-term) and 37 (13 long-term) studies reported on major and general AEs, respectively. Table H presents the key findings for general AEs in comparisons between different SGAs.

Major AEs. Over the long term, aripiprazole appears to increase the risk for developing diabetes compared with risperidone. One large retrospective review of a Medicaid database found that patients newly initiating antipsychotics (compared with propensity-score matched controls not on antipsychotics) were at higher risk ($p < 0.0001$) for developing diabetes after >1 year followup if taking aripiprazole (HR 7.72, 95% CI 3.70 to 16.12) compared with risperidone (HR 2.20, 95% CI 1.14 to 4.26). These results were inconsistent with another small long-term study of 47 patients on various SGAs that only found

one incidence of diabetes in a patient taking clozapine. Findings on other major AE outcomes were rated as insufficient SOE.

General AEs. To summarize the findings on general SAEs—

- **Body composition**. Risperidone probably decreases gains in weight (short-term) and BMI (short-and long-term) to a small extent compared with olanzapine; similar findings were found for quetiapine versus olanzapine over the long- but not short-term where there may be little or no difference. There appears to be little or no difference between weight gains caused by olanzapine and clozapine over short-term treatment. Quetiapine and risperidone are probably of little or no difference for short-term changes in BMI and 7 percent or greater increase in weight, and may be of little or no difference for BMI changes or weight gain over the long-term. For 7 percent or greater gain in body weight, there appears to be little or no difference between olanzapine and quetiapine, or olanzapine and risperidone.
- **Hyperprolactinemia**. Quetiapine may decrease the risk for hyperprolactinemia compared with risperidone.
- **Sedation**. There may be little or no difference between olanzapine and risperidone for risk of sedation.

All findings for clozapine versus risperidone and aripiprazole versus risperidone, and most findings for clozapine versus olanzapine, were rated as insufficient SOE, mainly due to imprecision but also because of risk of bias and inconsistency.

Table H. Summary of findings for general adverse effects: Short- and long-term findings of comparisons between different SGAs

Comparison (G1 vs. G2), Timeframe	Outcome	N Studies, N Patients	G1 Events	G1 N	G2 Events	G2 N	Relative Effects[a]	Strength of Evidence, Conclusions
Clozapine vs. Olanzapine Short-term	Weight (kg)	5 (136)	-	62	-	74	MD, -1.56; 95% CrI, -5.12 to 1.57	Low; may make little or no difference[b]
Olanzapine vs. Quetiapine Short-term	Weight (kg)	3 (232)	-	116	-	116	MD, 4.00; 95% CrI, -1.67 to 10.79	Low; may make little or no difference[c]
	BMI ($kg.m^{-2}$)	3 (232)	-	116	-	116	MD, 1.36; 95% CrI, -0.29 to 3.40	Low; may make little or no difference[c]
	≥ 7% increase in weight	3 (192)	72	99	47	93	RR: 1.41; 95% CI, 0.65 to 2.83	Low; may make little or no difference[c]
Olanzapine vs. Quetiapine Long-term	Weight (kg), 6 to <12months	3 (185)	-	90	-	95	MD, 7.91; 95% CrI, 3.65 to 12.29	Moderate; Quetiapine probably better[d]
	BMI ($kg.m^{-2}$), 6 to <12months	4 (203)	-	99	-	104	MD, 2.68; 95% CrI, 0.96 to 4.27	Moderate; Quetiapine probably better[d]
Olanzapine vs. Risperidone Short-term	Weight (kg)	13 (936)	-	331	-	605	MD, 2.18; 95% CrI, 1.13 to 3.25	Moderate; Risperidone probably slightly better[d]
	BMI ($kg.m^{-2}$)	9 (737)	-	244	-	493	MD, 0.94; 95% CrI, 0.64 to 1.30	Moderate; Risperidone probably slightly better[d]
Olanzapine vs. Risperidone Short-term (continued)	≥ 7% increase in weight	6 (504)	107	150	188	354	RR, 1.36; 95% CrI, 0.93 to 2.04	Low; may make little or no difference[c]
	Sedation	7 (321)	35	133	36	188	RR, 1.19; 95% CrI, 0.73 to 2.35	Low; may make little or no difference[c]

Table H. Summary of findings for general adverse effects: Short- and long-term findings of comparisons between different SGAs (continued)

Comparison (G1 vs. G2), Timeframe	Outcome	N Studies, N Patients	G1 Events	G1 N	G2 Events	G2 N	Relative Effects[a]	Strength of Evidence, Conclusions
Olanzapine vs. Risperidone Long-term	Weight (kg), 6 to <12months	4 (295)	-	85	-	210	MD, 4.40; 95% CrI, -0.54 to 9.86	Low; may make little or no difference[c]
	BMI (kg.m-2), 6 to <12months	5 (328)	-	94	-	234	MD, 1.66; 95% CrI, 0.19 to 3.42	Moderate; Risperidone probably slightly better[d]
	≥ 7% increase in weight, 6 to <12 months	3 (264)	28	64	64	200	RR: 1.44; 95% CI, 0.55 to 5.50}	Low; may make little or no difference[c]
Quetiapine vs. Risperidone Short-term	Weight (kg)	3 463)	-	116	-	347	MD, 0.08; 95% CrI, -3.77 to 3.14	Low; may make little or no difference[f]
	BMI (kg.m-2)	3 (463)	-	116	-	347	MD, 0.04; 95% CrI, -1.34 to 1.20	Moderate; probably makes little or no difference[d]
	≥ 7% increase in weight	4 (417)	55	104	176	313	RR: 0.91; 95% CI, 0.56 to 1.44	Moderate; probably makes little or no difference[d]
	Hyper-prolactinemia	4 (118)	4	31	45	87	RR, 0.20; 95% CrI, 0.06 to 0.73	Low; Quetiapine may decrease risk[e]
Quetiapine vs. Risperidone Long-term	Weight (kg), 6 to <12months	3 (295)	-	93	-	202	MD, -1.48; 95% Crl, -4.16 to 1.18	Low; may make little or no difference[e]
	BMI ($kg.m^{-2}$), 6 to <12months	4 (328)	-	102	-	226	MD, -0.32; 95% CrI, -1.56 to 1.12	Low; may make little or no difference[e]

BMI=body mass index; CrI = credible interval; kg = kilogram; m = meters; MD = mean difference; N=number; RR = risk ratio

[a] Positive MDs favor group 2; RR above 1.0 favor group 2

[b] Downgraded for ROB and imprecision, because CrI includes possibility for clinically relevant benefit for group 1.

[c] Downgraded for ROB and imprecision, because CrI includes possibility for clinically relevant benefit for group 2.

[d] Downgraded for ROB.

[e] Downgraded for ROB and imprecision, based on small sample size.

[f] Downgraded for ROB and inconsistency.

SGAs Versus SGAs: Dose Comparisons

The effects between different doses of SGAs in terms of major AEs during short-term treatment are mostly unknown (insufficient SOE). There may be no difference between 5 mg/day and 10 mg/day asenapine for risk of developing diabetes over 8 weeks of treatment (low SOE); both groups (n = 98, n = 102) had 7 percent incidence of possible new-onset diabetes (compared with 4% in placebo group).

Table I includes the findings for general AEs; the doses considered are identified for each drug. The findings for each drug are summarized below.

- **Aripiprazole**. Different doses of aripiprazole are probably of little or no difference in the extent of weight gain they cause over the short-term. There may be little or no difference between doses for any EPS symptoms, BMI, the proportion gaining 7 percent or more weight, and somnolence (all short-term); for these

outcomes the 95% CIs included values favoring the low dose. There appears to be little or no difference in risk for hypertriglyceridemia or high total cholesterol.

- **Asenapine**. There is probably little or no difference in the short-term between low and high doses of asenapine for weight gain, proportion of patients gaining 7 percent or more weight, risk of somnolence, or risk of hyperprolactinemia.
- **Quetiapine**. Low and high doses of quetiapine are likely of little or no difference for risk of gaining greater than 7 percent weight, somnolence, or sedation over the short-term.
- **Risperidone**. Risks for somnolence and EPS symptoms may be of little or no difference for low- versus high-dose risperidone during short-term treatment.

Table I. Summary of findings for general adverse effects: Short-term findings from comparisons between different doses of SGAs

Comparison	Outcome	High Dose Events	High Dose N	Low Dose Events	Low Dose N	Relative Effects[a]	Strength of Evidence; Conclusions
Aripiprazole High (15/30mg/day) vs. Low (10mg/day)	Any EPS	39 12	99 54	23 13	98 59	RR, 1.68; 95% CI, 1.09 to 2.59 RR, 1.01; 95% CI, 0.50 to 2.02	Low; may make little or no difference[b]
	Weight (kg)	-	229	-	234	MD, 0.22; 95% CrI, -0.64 to 1.09	Moderate; probably makes little or no difference[c]
	BMI ($kg \cdot m^{-2}$)	-	223	-	233	MD, 0.14; 95% CrI, -0.47 to 5.86	Low; may make little or no difference[b]
	≥ 7% weight increase	37	250	24	256	RR, 1.62; 95% CrI, 0.47 to 5.86	Low; may make little or no difference[b]
	High cholesterol	28 0	65 54	27 0	64 59	RR, 1.02; 95% CI, 0.68 to 1.52 Not estimable	Low; may make little or no difference[d]
	High triglycerides	22 2	65 54	22 6	65 59	RR, 1.00; 95% CI, 0.62 to 1.62 RR: 0.36; 95% CI, 0.08 to 1.73	Low; may make little or no difference[d]
	Somnolence	62	255	47	257	RR, 1.31; 95% CrI, 0.46 to 3.80	Low; may make little or no difference[b]
Asenapine High (10mg/day) vs. Low (5mg/day)	BMI ($kg \cdot m^{-2}$)	--	-	-	-	MD, 0.03; 95% CI, -0.04 to 0.10	Low; may make little or no difference[e]
	≥ 7% weight increase	10 8	99 90	9 11	95 92	RR, 1.07; 95% CI, 0.45 to 2.51 RR, 0.74; 95% CI, 0.31 to 1.76	Moderate; probably makes little or no difference[e]
	Somnolence	31 52	106 99	24 49	98 104	RR, 1.19; 95% CI, 0.76 to 1.89 RR, 1.11; 95% CI, 0.85 to 1.47	Moderate; probably makes little or no difference[e]
	Hyperprolactinemia	20	106	23	98	RR, 1.24; 95% CI, 0.73 to 2.12	Low; may make little or no difference[e]
Quetiapine High (600/800 mg/day) vs. Low (400 mg/day)	≥ 7% weight increase	14 10	74 98	17 14	73 95	RR, 0.81; 95% CI, 0.43 to 1.52 RR, 0.69; 95% CI, 0.32 to 1.48	Moderate; probably makes little or no difference[c]
	Somnolence	22 31	74 98	20 27	73 95	RR, 1.09; 95% CI, 0.65 to 1.81 RR, 1.11; 95% CI, 0.72 to 1.71	Moderate; probably makes little or no difference[c]
	Sedation	4 25	74 98	4 22	73 95	RR, 0.99; 95% CI, 0.26 to 3.80 RR, 1.10; 95% CI, 0.67 to 1.81	Moderate; probably makes little or no difference[c]

Table I. Summary of findings for general adverse effects: Short-term findings from comparisons between different doses of SGAs (continued)

Comparison	Outcome	High Dose Events	High Dose N	Low Dose Events	Low Dose N	Relative Effects[a]	Strength of Evidence; Conclusions
Risperidone High (3- 6mg/ day) vs. Low (0.5-3mg/day)	Any EPS	20 15	51 61	18 4	55 50	RR, 1.20; 95% CI, 0.72 to 2.00 RR, 3.07; 95% CI, 1.09 to 8.68	Low; may make little or no difference[b]
	Somnolence	6 34	51 61	13 21	55 50	RR, 0.50; 95% CI, 0.20 to 1.21 RR, 1.33; 95% CI, 0.89 to 1.97	Low; may makie little or no difference[f]

AE = adverse effect; BMI=body mass index; CI = confidence interval; CrI = credible interval; EPS = extrapyramidal symptoms; kg = kilogram; m = meter; mg = milligrams; MD = mean difference; N=number; RR = risk ratio

[a] Positive MDs and RRs above 1.0 favor the low dose group. Effects are shown for each study contributing data (we did not pool data from only 2 studies).

[b] Downgraded for ROB and imprecision, because CIs include possibility for clinically relevant benefit for the low dose group.

[c] Downgraded for ROB.

[d] Downgraded for ROB and imprecision due to small sample sizes.

[e] Downgraded for imprecision, because CIs include possibility for clinically relevant benefit for the low dose group.

[f] Downgraded for ROB and imprecision, because of inconsistency between studies.

FGAs Versus Placebo

No findings for major or general AEs in comparisons between FGAs and placebo offered greater than insufficient SOE. Four small studies reported on AEs to a varying extent with most outcomes having imprecise data from one small study having medium or higher ROB.

SGAs Versus Placebo

Findings for major and general AEs in comparisons between SGAs and placebo are presented below.

Major AEs

There is probably little or no difference in the short-term across all SGAs compared with placebo for mortality (13 studies, 2447 patients; 0 events) or for having a pathologically prolonged QT interval (14 studies, 2425 patients; events in 19 of 1490 in SGA and 9 of 935 in placebo).

Compared with no antipsychotic treatment, SGAs may increase the risk for developing diabetes over the long-term. A large retrospective cohort study compared incidence of type 2 diabetes in patients newly initiated on antipsychotics compared with matched patients not taking antipsychotics for at least 1 year; taking SGAs was associated with an increased risk (HR 2.89, 95% CI 1.64 to 5.10; 25.3 vs. 7.8 cases per 10,000 person-years followup).

Other outcomes were rated as having insufficient SOE due to rare events (≤ 5% of patients) occurring in samples too small to offer adequate power to detect a difference (N < 2000).

General AEs

Tables J and K summarize findings for general AEs having at least low SOE during short- and long-term studies, respectively. A summary of the key points is included below for findings across SGAs and for individual drugs, respectively.

- All SGAs versus placebo. SGAs as a class are probably worse than placebo/no antipsychotic treatment for seven outcomes: EPS symptoms, changes to body composition (weight, BMI, and ≥7% weight gain), high triglycerides, sedation, and somnolence. They appear to be worse for risk of high total cholesterol, and there may be little or no difference in risk for akathisia. In the longer term, few studies provided insufficient SOE.
- Individual SGAs versus placebo.
  - Aripiprazole is probably slightly worse than placebo/no treatment for gains in weight and BMI, and may increase risk for any EPS, ≥7 percent weight gain, and somnolence.

- Compared with placebo, olanzapine likely increases weight gain and BMI, and may increase risk for ≥7 percent weight gain and hyperprolactinemia.
- Quetiapine probably increases weight gain slightly, and may make little or no difference in risk for sedation and somnolence.
- Risperidone probably increases weight gain and BMI to a small extent, and probably increases risk for somnolence. It may increase risk for any EPS symptoms. In long-term studies, there may be little or no difference over placebo in changes in weight and BMI.
- Ziprasidone probably makes little or no difference for weight gain, and appears to make little or no difference for somnolence.

Table J. Summary of findings for general adverse effects: Short- and long-term durations of comparisons between SGAs and placebo

Comparison	Outcome	N Studies, N Patients	SGA Events	SGA N	Placebo Events	Placebo N	Relative Effects[a]	Strength of Evidence; Conclusions
All SGAs vs. placebo	Any EPS	15, 2730 2, 32	233 0	1757 17	40 0	973 15	RR, 2.94; 95% CI, 2.02 to 4.27 Not estimable	Moderate; SGAs probably increase risk[b]
	Akathisia	21, 3638	151	2433	56	1205	RR, 1.29; 95% CrI, 0.81 to 2.27	Low; SGAs may make little or no difference[c]
	Weight (kg)	37, 3919	-	2384	-	1535	MD, 1.48; 95% CI, 1.06 to 1.91	Moderate; SGAs probably increase slightly[b]
	BMI ($kg.m^{-2}$)	16, 2462	-	1582	-	880	MD, 0.61; 95% CI, 0.38 to 0.85	Moderate; SGAs probably increase slightly[b]
	≥ 7% increase in weight	17, 3057	337	2023	42	1034	RR, 3.53; 95% CrI, 2.49 to 5.23	Moderate; SGAs probably increase risk[b]
	Increased total cholesterol	6, 643 1, 218	92 0	410 52	13 0	233 166	RR, 3.17; 95% CrI, 1.29 to 9.13 Not estimable	Low; SGAs may increase risk[d]
	Increased triglycerides	10, 1383	130	897	38	486	RR, 1.64; 95% CrI, 1.09 to 2.63	Moderate; SGAs probably increase risk[b]
	Sedation	21, 2710	288	1696	79	1014	RR, 2.19; 95% CrI, 1.50 to 3.41	Moderate; SGAs probably increase risk[b]
	Somnolence	26, 3942	560	2481	119	1461	RR, 2.91; 95% CrI, 2.27 to 3.86	Moderate; SGAs probably increase risk[b]

Table J. Summary of findings for general adverse effects: Short- and long-term durations of comparisons between SGAs and placebo (continued)

Comparison	Outcome	N Studies, N Patients	SGA Events	SGA N	Placebo Events	Placebo N	Relative Effects[a]	Strength of Evidence; Conclusions
Aripiprazole vs. placebo	Any EPS	6, 1000	117	655	17	345	RR, 3.10; 95% CrI, 1.26 to 7.01	Low; Aripiprazole may increase risk[e]
	Weight (kg)	7, 1042	-	647	-	395	MD, 0.98; 95% Crl, 0.54 to 1.48	Moderate; Aripiprazole probably increases slightly[b]
	BMI ($kg.m^{-2}$)	5, 881	-	587	-	294	MD, 0.33; 95% CI, 0.07 to 0.67	Moderate; Aripiprazole probably increases slightly[b]
	≥ 7% increase in weight	5, 991	93	647	15	344	RR, 3.01; 95% Crl, 1.33 to 7.10	Low; Aripiprazole may increase risk[e]
	Somnolence	6, 1012	119	661	29	351	RR, 2.73; 95% Crl, 1.24 to 7.65	Low; Aripiprazole may increase risk[e]
Olanzapine vs. placebo	Weight (kg)	4, 337	-	215	-	122	MD, 3.96; 95% CI, 2.31 to 6.34	Moderate; Olanzapine probably increases[b]
	BMI ($kg.m^{-2}$)	2, 267	- -	107 72	- -	54 34	MD, 1.16; 95% CI, 0.93 to 1.39 MD, 1.50; 95% CI, 1.06 to 1.94	Moderate; Olanzapine probably increases[b]
	≥ 7% increase in weight	4, 337	99	215	8	122	RR, 6.08; 95% Crl, 1.84 to 27.06	Low; Olanzapine may increase risk[e]
	Hyper-prolactinemia	2, 268	50 58	107 72	1 6	54 35	RR, 25.53; 95% CI, 3.58 to 177.76 RR, 4.70; 95% CI, 2.25 to 9.82	Low; Olanzapine may increase risk[e]
Quetiapine vs. placebo	Weight (kg)	6, 778	-	473	-	305	MD, 1.44; 95% CI, 0.60 to 2.31	Moderate; Quetiapine probably increases slightly[b]
	Sedation	6, 778	90	473	32	305	RR, 1.67; 95% Crl, 0.77 to 3.87	Low; may make little or no difference[c]
	Somnolence	3, 697	106	432	18	265	RR, 2.95; 95% Crl, 0.92 to 8.62	Low; may make little or no difference[c]

Table J. Summary of findings for general adverse effects: Short- and long-term durations of comparisons between SGAs and placebo (continued)

Comparison	Outcome	N Studies, N Patients	SGA Events	SGA N	Placebo Events	Placebo N	Relative Effects[a]	Strength of Evidence; Conclusions
Risperidone vs. placebo	Any EPS	5, 636	52	365	13	271	RR, 2.78; 95% CrI, 1.27 to 6.50	Low; Risperidone may increase risk[e]
	Weight (kg)	14, 929	-	522	-	475	MD, 1.52; 95% CI, 0.78 to 2.29	Moderate; Risperidone probably increases slightly[b]
	BMI ($kg.m^{-2}$)	6, 730	-	397	-	333	MD, 0.68; 95% CI, 0.27 to 1.18	Moderate; Risperidone probably increases slightly[b]
	Somnolence	9, 862	163	473	43	389	RR, 3.25; 95% Crl, 1.96 to 5.94	Moderate; Risperidone probably increases risk[b]
Ziprasidone vs. placebo	Weight (kg)	3, 360	-	246	-	114	MD, -0.10; 95% CI, -1.34 to 1.13	Moderate; Ziprasidone probably makes little or no difference[b]
	Somnolence	3, 548	76	358	13	190	RR, 2.97; 95% Crl, 0.84 to 9.96	Low; Ziprasidone may make little or no difference[c]

AE = adverse effect; BMI = body mass index; CI = confidence interval; CrI = credible interval; kg = kilogram; m = meter; MD = mean difference; N = number; RR = risk ratio; SGA = second-generation antipsychotic

[a] Risk ratios above 1.0 and positive MD favor placebo.

[b] Downgraded for ROB.

[c] Downgraded for ROB and imprecision because point estimate and CrI includes clinically significant favor for placebo.

[d] Downgraded for ROB and inconsistency.

[e] Downgraded for ROB and imprecision, based on small sample size.

Table K. Summary of findings for general adverse effects: Long-term durations of SGAs versus placebo

Comparison	Outcome, Duration	N Studies, N Patients	Relative Effects[a]	Strength of Evidence; Conclusions
Risperidone vs. placebo	Weight (kg), 6 to <12months	4, 467	MD, 2.86; 95% CrI, -1.22 to 7.42	Low; Risperidone may make little or no difference[b]
	BMI (kg.m-2), 6 to <12months	2, 405	MD, 0.70; 95% CI, 0.49 to 0.91 MD, 1.80; 95% CI, -0.61 to 4.21	Low; Risperidone may make little or no difference[b]

BMI = body mass index; CI = confidence interval; CrI = credible interval; kg = kilogram; m = meter; MD = mean difference; N = number; RR = risk ratio; SGA = second-generation antipsychotic

[a] Positive MD favors placebo.

[b] Downgraded for ROB and imprecision because CrI includes clinically significant favor for placebo.

Between- and Within-Study Subgroup Effects

Bayesian univariate meta-regression analyses were conducted to determine if effects on four outcomes (weight change, proportion gaining 7% or more weight, somnolence, and EPS symptoms) were influenced by four subgroup variables (mean age, % male, % treatment naïve, and treatment duration). We used data from longest followup duration from SGA-placebo/no treatment comparisons. For the outcome of EPS symptoms, we included data from findings on (in hierarchical order) akathisia, dystonia, and any EPS. The only analysis with statistically significant findings was for treatment duration on weight change; age and proportion being treatment naïve were not found to significantly modify effects. The model predicted small increments in weight gain over longer treatment durations (0.043 kg per week; 95% CrI, 0.015 to 0.072). Because of these findings, we ran adjusted network meta-analyses for weight and BMI using the study-level variable of treatment duration; athough this variable was shown to statistically modify effects, the results of the network meta-analysis were not changed to any meaningful extent.

Observations based on diagnostic condition did not indicate any moderating effect in terms of the four harm outcomes evaluated; harms appeared to occur to a similar magnitude in different conditions regardless of the typical dose used.

Twenty-six studies reported on subgroup analyses. Findings were often inconsistent on whether there are any moderating effects by various subgroup variables on harms. Several studies found no significant differences in harms for different age groups. Body composition, fasting glucose, and prolactin elevations do not appear to differ in patients taking SGAs based on concurrent use of psychostimulants. Dose of SGAs—particularly when considering cumulative doses—was found in two large observational studies to increase the risk for metabolic effects including increased glucose levels and development of diabetes. Risperidone appears to increase serum prolactin more in females than males; few studies reported on other subgroup variables for this harm. Findings for effect moderation on risk for somnolence and neuromotor effects were mainly from single studies.

Applicability of Findings

Study populations seem moderately applicable to general practice in terms of age, gender and existence of common comorbid diagnoses (e.g., ADHD comorbidity within primary diagnosis of bipolar or tic disorders) within each condition category. Findings will not be as applicable in terms of patients having complex clinical diagnoses, medical comorbidity, less-than-moderate symptom severity, and (with the exception of studies of clozapine in schizophrenia) a history of poor response to antipsychotics.

The majority of the studies in this review did not enroll young adults; therefore, the results may have limited applicability to this population. Nor was the mean age in any condition below 8 years. Exclusion of patients with comorbidities, a history of various adverse events, and/or less-than-moderate symptom severity at baseline may have overestimated the estimates of the efficacy and underestimated the harms of antipsychotics.

Another factor that restricts the applicability of the studies is the short duration of followup (75% of studies had treatment durations < 6 months). Adequate

trials of antipsychotic treatment to assess response can be considered within 4 to 6 weeks,[16] which supports applicability for these outcomes from the evaluated studies; nevertheless, issues impacting longterm treatment success, such as treatment compliance and resistance, were not accounted for in many studies. Data on most effectiveness outcomes were deficient, and few studies allowed for conclusions on major adverse effects ~~es~~especially those often arising with longterm treatment (e.g., tardive dyskinesias, diabetes). Adverse effects may have been underestimated due to the short followup periods; not all effects are likely to become evident in all patients within the 1-2 month treatment phase commonly investigated.

Applicability may also be limited due to monitoring practices within the trial settings to ensure treatment adherence as well as perform dose adjustments based on response and tolerability assessments. In typical practice settings, it is likely that will patients have lower rates of medication adherence—and therefore less symptom improvement—and may have higher rates of AEs because of poor monitoring. Although comprehensive and individualized monitoring for AEs has been recommended for several years,[12,16,29] there is evidence from Medicaid claims data[30-32] and clinician self-reports[33] that these practices remain inadequate. Guidelines for screening and monitoring have been developed, especially in the area of schizophrenia where antipsychotics are the primary treatment, although there has been some critique of their degree of rigor (e.g., use of systematic reviews of the evidence), stakeholder involvement, and efforts to make recommendations on organizational aspects.[34]

Implications for Clinical and Policy Decisionmakers

There are some conclusions which can support clinician decisionmaking despite at best moderate SOE. SGAs showed benefit over placebo for manic and mixed states in bipolar disorder, irritability and other symptoms in autism, and aggression and conduct problems in children with DICD with or without comorbid ADHD. It is not known whether antipsychotics improve clinical impressions of severity and hyperactivity in youth who have previously responded to psychostimulant medications. Moderate evidence for clinical benefit in these symptoms is present only for those for whom stimulant medications have not produced clinically significant reductions in ADHD symptoms, or for whom DICD is the primary diagnosis. Interestingly, comorbid ADHD did not impact the treatment effect across many conditions, and there was a significant placebo effect for treatment of positive and negative symptoms of schizophrenia. Limited evidence suggests that SGAs are effective for reduction in tic severity. The effect on depressive symptoms may be small and possibly nonsignificant for schizophrenia and bipolar disorder. Reliance on findings from placebo-controlled studies for schizophrenia may not offer great help to those needing to choose between different antipsychotics for this condition which often relies on this treatment. In general, the small number of comparions between different antipsychotics is a limitation in the evidence base. Some of the findings for harms are quite considerable in light of the short-term duration of treatment of many of the studies contributing data. Nevertheless, some findings on harms—such as the low impact on weight suggested by studies of molindone—may provide some assistance when choosing between treatment alternatives. Continued guidance related to ongoing benefit-harm assessments for individual patients, regardless of which antipsychotic is prescribed, seems prudent.

Consistent with the role of systematic reviewers, we did not incorporate contextual considerations in our assessment of the SOE as would guideline developers.[26] For example, our assessment of precision in findings should be interpreted in view of our confidence in the direction and magnitude of the average effect and an estimated threshold rather than having a (possibly greater) threshold based on various benefit-harm considerations. Several of the findings for intermediate outcomes only support small effects, although the placebo effect in several studies (especially for schizophrenia) was substantial which makes some findings difficult to interpret in light of real-world practice. Likewise, we did not downgrade any evidence for lack of directness related to the comparability of study populations with those treated in clinical practice, for which there may be important differences. The main reasons we downgraded the SOE was for risk of bias (largely from incomplete data due to study withdrawals) and imprecision from small samples or when the results included possibility of substantial benefit or harm when insignificant findings were found (i.e., limiting confidence in findings of no difference). It should be recognized that attaining high SOE from trials of antipsychotics in children with psychiatric conditions is likely very difficult and the overall evidence reviewed should not be interpreted as lacking in credibility.

Systematic reviews may become outdated, at least in part, if new studies are published that change some or all of their conclusions. Although our comprehensive search was only undertaken to April 2016, we are quite confident there

has been no evidence as of September 2016 which would change our findings in such a manner (e.g., to moderate or higher SOE for any outcome). A search update in Medline for April to September 8, 2016 identified three RCTs[35-37] and one retrospective cohort study;[38] assessment of these studies for their ability to potentially change the SOE indicated no change for the relevant comparators and outcomes. The studies, though, appear to represent a trend for more comparative research between different SGAs, if not also between SGAs and FGAs as suggested from our findings.

Research Gaps

The following general recommendations for future research are based on the preceding discussion regarding the limitations of the current evidence:

- Studies examining long-term effectiveness and, particularly, safety of antipsychotics (and differences between different antipsychotics) over the course of several years are needed. Future research should evaluate long-term developmental outcomes, such as growth, maturation, and cognitive and emotional development.
- Future studies should evaluate outcomes that are important to patients and parents, including health-related quality of life, school performance, and involvement with the legal system.
- Studies examining the impact of key patient subpopulations on important outcomes are needed to inform clinical practice. In particular, subgroup analyses examining young adults would be helpful in guiding clinical decisions due to the unique issues associated with this population.
- Consensus on outcomes and outcome measures is needed to ensure consistency and comparability across future studies. Moreover, consensus on minimal clinically important differences is needed to guide study design and interpretation of results.
- Large-scale effectiveness studies that use inclusive patient-selection criteria and closely match typical clinical practice are needed to achieve greater applicability of results. Data on the real-world benefits and harms across groups defined by race/ethnicity, socioeconomic status, and geographical region would be informative.
- Studies incorporating therapeutic drug monitoring over long-term periods in naturalistic settings should be encouraged to help create quality standards and provide insight into operational considerations to inform recommendations for monitoring.
- Considering antipsychotics are recommended for use as adjunctive, or add-on, treatment for many conditions/symptoms, more studies examining these approaches (e.g., behavioral/family interventions with and without antipsychotics for hyperactivity or irritability) may help practitioners create guidance on when to start a trial of antipsychotics

Conclusions

The efficacy and safety of FGAs and SGAs have been studied in children, adolescents, and young adults (ages ≤ 24 years) for a wide array of psychiatric conditions. Overall, data for head-to-head comparisons (FGAs vs. SGAs, FGAs vs. FGAs, and SGAs vs. SGAs) were generally of insufficient or low SOE; therefore, few conclusions regarding the relative benefits and harms of antipsychotics could be drawn. For schizophrenia, there appears to be little or no difference between FGAs and SGAs for negative symptoms, positive symptoms, response rates, and global impressions of illness severity; deciding on which antipsychotic to use for this condition likely relies on close examination of the relative harms including considerations of their tolerance, management, and reversibility. Many conclusions for intermediate outcomes of SGAs relative to placebo showed small magnitudes of effect, and this together with some confidence that SGAs increase the risk for several adverse effects with potentially long-term health consequences lends towards a fine balance of benefits and harms particularly in cases where alternatives exist. Evidence was sparse for several patient- and family-important outcomes, such as health-related quality of life, involvement with the legal system, and school performance. Our confidence in the findings from studies reporting most long-term data was poor.

Treatment benefit and harms were examined most frequently for schizophrenia. Fewer studies examined other conditions; only one study was eligible for each of depression and obsessive-compulsive disorder, and there were no eligible studies exclusively examining posttraumatic stress disorder, anxiety disorders, or substance use disorder. Young adults were rarely examined, particularly for conditions other than schizophrenia; there were also few studies of young children. Additional research is needed to assess the treatment efficacy, and particularly the harms, of antipsychotics in these populations.

This review identified several areas for which the evidence is sparse and which are priorities for future research. One of the greatest priorities for future research is the systematic evaluation of harms. Studies incorporating therapeutic drug monitoring over long-term periods in naturalistic settings could help create a more accurate picture of the comparative harms between the diverse number of antipsychotics. They may also help define quality standards and provide insight into operational considerations to inform recommendations for monitoring implementation. Comprehensive comparative effectiveness reviews such as this one, combined with active involvement of patients, families, and multidisciplinary practitioners may improve the applicability and usefulness of guidelines and help ensure their recommendations can be attained.

References

1. Olfson M, Crystal S, Huang C, et al. Trends in antipsychotic drug use by very young, privately insured children. J Am Acad Child Adolesc Psychiatry. 2010 Jan;49(1):13-23. PMID: 20215922.
2. Pathak P, West D, Martin BC, et al. Evidence-based use of second-generation antipsychotics in a state Medicaid pediatric population, 2001-2005. Psychiatr Serv. 2010 Feb;61(2): 123-9. doi: 10.1176/appi.ps.61.2.123. PMID: 20123816.
3. Zito JM, Safer DJ, de Jong-van den Berg LT, et al. A three-country comparison of psychotropic medication prevalence in youth. Child Adolesc Psychiatry Ment Health. 2008;2(1):26. doi: 10.1186/1753-2000-2-26. PMID: 18817536.
4. Zito JM, Safer DJ, Sai D, et al. Psychotropic medication patterns among youth in foster care. Pediatrics. 2008 Jan;121(1): e157-63. doi: 10.1542/peds.2007-0212. PMID: 18166534.
5. Olfson M, King M, Schoenbaum M. Treatment of young people with antipsychotic medications in the United States. JAMA Psychiatry. 2015 Sep;72(9):867-74. doi: 10.1001/jamapsychiatry.2015.0500. PMID: 26132724.
6. Zito JM, Derivan AT, Kratochvil CJ, et al. Off-label psychopharmacologic prescribing for children: history supports close clinical monitoring. Child Adolesc Psychiatry Ment Health. 2008;2(1):24. doi: 10.1186/1753-2000-2-24. PMID: 18793403.
7. 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.
8. 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. Eur Neuropsychopharmacol. 2011 Aug;21(8):621-45. doi: 10.1016/j.euroneuro.2010.07.002. PMID: 20702068.
9. 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 Psychiatry. 2010 May;19(2):124-37. PMID: 20467549.
10. 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 Saf. 2011 Aug 1;34(8):651-68. doi: 10.2165/11592020-000000000-00000. PMID: 21751826.
11. Correll CU, Penzner JB, Parikh UH, et al. Recognizing and monitoring adverse events of second-generation antipsychotics in children and adolescents. Child Adolesc Psychiatr Clin N Am. 2006 Jan;15(1):177-206. doi: 10.1016/j.chc.2005.08.007. PMID: 16321730.
12. Rettew DC, Greenblatt J, Kamon J, et al. Antipsychotic medication prescribing in children enrolled in Medicaid. Pediatrics. 2015 Apr;135(4):658-65. doi: 10.1542/peds.2014-2260. PMID: 25733747.
13. Birnbaum ML, Saito E, Gerhard T, et al. Pharmacoepidemiology of antipsychotic use in youth with ADHD: trends and clinical implications. Curr Psychiatry Rep. 2013 Aug;15(8): 382. doi: 10.1007/s11920-013-0382-3. PMID: 23881713.
14. Crystal S, Olfson M, Huang C, et al. Broadened use of atypical antipsychotics: safety, effectiveness, and policy challenges. Health Aff (Millwood). 2009 Sep-Oct;28(5):w770-81. doi: 10.1377/hlthaff.28.5.w770. PMID: 19622537.
15. Schmid I, Burcu M, Zito JM. Medicaid prior authorization policies for pediatric use of antipsychotic medications. JAMA. 2015 Mar 3;313(9):966-8. doi: 10.1001/jama.2015.0763. PMID: 25734740.
16. Findling RL, Drury SS, Jensen PS; American Academy of Child and Adolescent Psychiatry. Practice Parameter for the use of Atypical Antipsychotic Medications in Children and Adolescents. 2012. Available at www.aacap.org
17. Seida JC, Schouten JR, Mousavi SS, et al. First- and Second-Generation Antipsychotics for Children and Young Adults. Comparative Effectiveness Review No. 39. AHRQ Publication No. 11(12)-EHC077-EF. Rockville MD: Agency for Healthcare Research and Quality; 2012. Available at www.effectivehealthcare.ahrq.gov
18. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: AHRQ; 2014. Available at www.effectivehealthcare.ahrq.gov
19. First and Second Generation Antipsychotics for Children and Young Adults-Comparative Effectiveness Review Update Protocol Agency of Healthcare Research and Quality. 2015. https://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2149
20. 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

21. Wells G, Shea B, O'Connell N, et al. The Newcastle-Ottawa Scale for Assessing the Quality of Nonrandomized Studies in Meta-analyses. Ottawa, ON: Department of Epidemiology and Community Medicine, University of Ottawa 2009.

22. 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). doi: 10.7326/m13-2886. PMID: 24727843.

23. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res. 2001 Aug;10(4):277-303. doi: 10.1177/096228010101000404. PMID: 11491414.

24. Golder S, Loke YK, Bland M. Meta-analysis of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. PLoS Med. 2011;8(5): e1001026. doi: 10.1371/journal.pmed.1001026. PMID: 21559325.

25. 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. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: AHRQ January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

26. 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. doi: 10.1016/j.jclinepi.2011.01.012. PMID: 21839614.

27. Glenton C, Santesso N, Rosenbaum S, et al. Presenting the results of Cochrane systematic reviews to a consumer audience: a qualitative study. Med Decis Making. 2010; 30:566-577. doi: 10.1177/0272989X10375853. PMID: 20643912

28. Storch EA, De Nadai AS, Lewin AB, et al. Defining treatment response in pediatric tic disorders: A signal detection analysis of the Yale Global Tic Severity Scale. J Child Adolesc Psychopharmacol. 2011 01 Dec;21(6):621-7. doi: http://dx.doi.org/10.1089/cap.2010.0149. PMID: 2012000755.

29. American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, North American Association for the Study of Obesity. Consensus development conference on antipsychotic drugs and obesity and diabetes. J Clin Psychiatry. 2004 Feb;65(2):267-72. PMID: 15003083.

30. Edelsohn GA, Parthasarathy M, Terhorst L, et al. Measurement of metabolic monitoring in youth and adult Medicaid recipients prescribed antipsychotics. J Manag Care Spec Pharm. 2015 Sep;21(9):769-77, 77a-77cc. doi: 10.18553/jmcp.2015.21.9.769. PMID: 26308224.

31. Morrato EH, Druss BG, Hartung DM, et al. Small area variation and geographic and patient-specific determinants of metabolic testing in antipsychotic users. Pharmacoepidemiol Drug Saf. 2011 Jan;20(1):66-75. doi: 10.1002/pds.2062. PMID: 21182154.

32. Morrato EH, Druss B, Hartung DM, et al. Metabolic testing rates in 3 state Medicaid programs after FDA warnings and ADA/APA recommendations for second-generation antipsychotic drugs. Arch Gen Psychiatry. 2010 Jan;67(1):17-24. doi: 10.1001/archgenpsychiatry.2009.179. PMID: 20048219.

33. Rodday AM, Parsons SK, Mankiw C, et al. Child and adolescent psychiatrists' reported monitoring behaviors for second-generation antipsychotics. J Child Adolesc Psychopharmacol. 2015 May;25(4):351-61. doi: 10.1089/cap.2014.0156. PMID: 25918843.

34. De Hert M, Vancampfort D, Correll CU, et al. Guidelines for screening and monitoring of cardiometabolic risk in schizophrenia: systematic evaluation. Br J Psychiatry. 2011 Aug;199(2):99-105. doi: 10.1192/bjp.bp.110.084665. PMID: 21804146

35. Ghanizadeh A. Twice-weekly aripiprazole for treating children and adolescents with tic disorder, a randomized controlled clinical trial. Ann Gen Psychiatry. 2016;15:21. PMID: 27579050.

36. Safavi P, Hasanpour-Dehkordi A, AmirAhmadi M. Comparison of risperidone and aripiprazole in the treatment of preschool children with disruptive behavior disorder and attention deficit-hyperactivity disorder: A randomized clinical trial. J Adv Pharm Technol Res. 2016;7:43-7. PMID: 27144151.

37. Lamberti M, Siracusano R, Italiano D, et al. Head-to-head comparison of aripiprazole and risperidone in the treatment of ADHD symptoms in children with autistic spectrum disorder and ADHD: a pilot, open-label, randomized controlled study. Pediatr Drugs. 2016;18: 319-328. PMID: 27278054.

38. Yoon Y, Wink LK, Pedapati EV, et al. Weight gain effects of second-generation antipsychotic treatment in autism spectrum disorder. J Child Adolesc Psychopharmacol. 2016 DOI: 10.1089/cap.2016.0049. PMID: 27389348.

Full Report

This executive summary is part of the following document: Pillay J, Boylan K, Carrey N, Newton A, Vandermeer B, Nuspl M, MacGregor T, Ahmed Jafri SH, Featherstone R, Hartling L. First- and Second-Generation Antipsychotics in Children and Young Adults: Systematic Review Update. Comparative Effectiveness Review No. 184. (Prepared by the University of Alberta Evidence-based Practice Center under Contract No. 290-2015-00001-I.) AHRQ Publication No. 17-EHC001-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2017. Errata January 2018. www.effectivehealthcare.ahrq.gov/reports/final/cfm. doi:10.23970/AHRQEPCCER184.

AHRQ Pub. No.17-EHC001-EF-1
March 2017
Errata January 2018
www.ahrq.gov