Evidence-based Practice Center Systematic Review Protocol

Project Title: *Treatments for Adults with Schizophrenia: A Systematic Review*

Initial publication date: August 17, 2016

Amendment Date(s): September 23, 2016

(Amendments Details–see Section VII)

## I. Background and Objectives for the Systematic Review

Schizophrenia is a serious mental health illness with potentially devastating effects on both patients and families. As a lifelong illness that frequently presents during early adulthood, it is associated with negative outcomes throughout the lifespan of affected individuals. The most recent version of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, version 5 (DSM-5),[1] has continued the trend of clarifying and simplifying the diagnostic criteria for schizophrenia from DSM-III through DSM-IV, without changing the defined patient population among the editions.[2] Currently, DSM-5 defines schizophrenia as the presence of two or more of the five core symptoms (delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms), with at least one of the symptoms being delusions, hallucinations, or disorganized speech, with symptoms being present for at least six months. Lifetime prevalence is approximately 0.3 to 0.7 percent, with onset most commonly between late adolescence through the third decade.[3] While clear mechanisms of disease and therefore means of prevention remain lacking, active lines of research seek to form hypotheses that connect known risk factors (e.g., genetics, season of birth, urban environments) to measurable neurocognitive deficits (e.g., prefrontal cortical deficits) and treatments shown to be effective (e.g., dopamine antagonism). The differential diagnosis is broad, with affective psychosis during depression or bipolar disorder and substance abuse comprising the bulk of the differential, with multiple other conditions that can lead to psychotic symptoms that are ruled out by appropriate use of diagnostic criteria and assessment. Antipsychotic medications (primarily effective via dopaminergic antagonism) usually provide the foundation of treatment, but non-drug treatments are almost always necessary to

complement the benefit of antipsychotic medications. Both pharmacological and nonpharmacological treatments for schizophrenia can result in meaningful improvements in a variety of outcome areas, including psychiatric symptoms, functioning (e.g., employment, social), service utilization (e.g., hospitalization, crisis services), legal system involvement, quality of life, self-harm and aggressive behaviors, treatment engagement and retention, and co-occurring substance abuse. Ideally, improvements in symptoms translate to long-term clinically relevant positive changes in other outcome areas (e.g., functioning outcomes) with limited and manageable adverse effects (e.g., sedation). Likely consistent with limited understanding of the causes and best treatments of schizophrenia, prognosis remains poor, with nearly 80 percent of patients continuing to require varying forms of social support throughout their lives. While pivotal trials of antipsychotic efficacy are limited to measurement of symptom reduction that may be of unclear clinical relevance, measurement of other important clinical outcomes that reflect improvement in social and occupational functioning are necessary to provide patients and families with the best management options possible.

Historically, the wide array of antipsychotic drug treatments has had a mixed impact on long-term outcomes, such as the ability to have consistent employment, successful interpersonal relationships, and maintain independent living, and serious concerns about adverse effects (e.g., tardive dyskinesia, weight gain, and diabetes and dyslipidemia) for some treatments. A patient's ability to adhere to and persist with treatment long-term is crucial, and can vary by treatment and patient characteristics, among many factors. Older, first-generation antipsychotics (FGAs), such as haloperidol, have proven efficacy but adverse effects, such as extrapyramidal symptoms and in some cases tardive dyskinesia, often limit long-term adherence. Second-generation antipsychotics (SGAs), beginning with clozapine, were introduced as having equal or better efficacy and the hope of fewer extrapyramidal symptoms and lower risk of tardive dyskinesia. But SGAs also have potentially serious adverse effects (e.g., cardiovascular and endocrinologic adverse effects) that make their overall risk/benefit profile less clear-cut than anticipated.

Since the American Psychiatric Association's (APA) most recent guidelines for treating adults with schizophrenia in 2004,[4] and a focused update in 2009,[5] important developments have occurred both in the pharmacologic and non-pharmacologic treatment

interventions for schizophrenia. There are at least five new antipsychotics and five new long-acting injectable formulations approved in the United States since the 2009 APA guideline update, and as many nonpharmacological (e.g., psychological) treatments developed, refined or expanded in recent years.[5-7] Given the ever-evolving nature of schizophrenia treatments and their potential for meaningful benefits and significant harms, up-to-date guidance based on comparative evidence is needed to determine the best treatment approach for individual patients.

There are several challenges remaining in the evidence base on treating adults with schizophrenia. These include evidence on effectiveness of complex interventions such as combined pharmacotherapy and psychotherapy, long-term benefits (e.g., employment, social relationships) with competing interventions, using consistent definitions of these outcomes and valid ascertainment methods. As there is no known cure for schizophrenia, achieving these long-term outcomes is tied to patients' ability to continue with treatment over the long-term, even life-long. As such, differences among the interventions in the ability to continue treatment may be important in selecting treatments. Equally important in selecting among competing interventions for a specific patient is consideration of patient-level characteristics that may affect the outcomes (including age, duration of symptoms, severity, and other psychiatric or medical co-morbidities) across a fairly diverse group of interventions. For example, treatment of negative symptoms (e.g., diminished emotional response, lack of interest) may differ from treatment of positive symptoms (e.g., hallucinations and delusions).

Treatment considerations will vary across the lifespan, and the success of any given treatment or combination of treatments depends on the balance of benefit and harm. Treating patients aggressively early in the disease is thought to improve long-term outcomes although overly aggressive treatments, particularly drug treatments, may result in adverse effects. Additionally, substance abuse (e.g., tobacco, alcohol, illicit drugs) may begin early on for patients with schizophrenia, and in these patients successful treatment of schizophrenia involves consideration of this comorbidity as well.[8-11] In middle-age and older patients, the metabolic effects related to both the disease and drug treatments become concerning.[12] Epidemiologic studies have found an association with obesity, elevated lipids and shorter lifespan, and several antipsychotic drugs are known to

increase these risks.[13-19] Older patients may have increased risk of mortality,[20-23] and require antipsychotic dosing changes with age.[24,25] Across all age groups, primary goals include diminishing core illness symptoms and reducing relapses of acute psychosis; however, physical health comorbidities can present challenges requiring modification of the treatment plan.

This comparative effectiveness review will provide a comprehensive review of current evidence that can help the APA and other organizations prepare updated guidance on how to treat patients with schizophrenia. The review will synthesize evidence on the effectiveness and comparative effectiveness of pharmacologic and non-pharmacologic treatment strategies for treating patients with schizophrenia, and will also highlight areas of controversy and identify needs for future research on the management of schizophrenia.

## II. The Key Questions

The draft scope for this topic was posted for public comment from April 28 – May 18, 2016. In response to public comments, we clarified some of the included populations, interventions, and outcomes; but no substantial changes were made to the scope.

1a. What are the comparative benefits and harms of pharmacological treatments for adults with schizophrenia?
1b. How do the benefits and harms of pharmacological treatments for adults with schizophrenia vary by patient characteristics*?
2a. What are the benefits and harms of psychosocial and other non-pharmacological treatments for adults with schizophrenia?
2b. How do the benefits and harms of nonpharmacological treatments for adults with schizophrenia vary by patient characteristics*?

*Patient characteristics include age, sex, race, ethnicity, socioeconomic status, time since illness onset, prior treatment history, co-occurring psychiatric disorders, pregnancy, etc.

- Population(s):
  - Adults with a diagnosis of schizophrenia, including those with co-occurring substance use disorders, and including those experiencing a first episode of schizophrenia (including those with schizophreniform disorder)
    - Key Question 1:
      - At least 90 percent of patients must have been diagnosed with schizophrenia.

- For studies specifically on harms of antipsychotic drugs, populations can be mixed-diagnoses, as the harms are not population-specific

  - Key Question 2: 50 percent of patients must have been diagnosed with a schizophrenia spectrum disorder diagnosis (i.e., schizophrenia, schizoaffective disorder, or schizophreniform disorder)[26]

- Interventions:
  - Key Question 1: Antipsychotic Medications

**Key First-Generation Antipsychotic Drugs**

| |
|---|
| Fluphenazine (Prolixin®*, Permitil®) |
| Haloperidol (Haldol®*) |
| Perphenazine (Trilafon®) |

*Includes extended-release IM injection formulation

Note: Limited to the most commonly-used first-generation antipsychotic drugs today.

**Second-Generation Antipsychotic Drugs**

| |
|---|
| Aripiprazole (Abilify®, Aristada™*) |
| Asenapine (Saphris® ), |
| Brexpiprazole (Rexulti®) |
| Cariprazine (Vraylar™) |
| Clozapine (Clozaril®, Fazaclo® ODT, Versacloz™) |
| Iloperidone (Fanapt®) |
| Lurasidone (Latuda®) |
| Olanzapine (Zyprexa®), Olanzapine Pamoate (Zyprexa® Relprevv™*) |
| Paliperidone (Invega®, Invega® Sustenna®*, Invega Trinza™*) |
| Quetiapine (Seroquel®, Seroquel XR®) |
| Risperidone (Risperdal®, Risperdal® M-TAB® ODT, Risperdal® Consta®*) |
| Ziprasidone (Geodon® ) |

*Includes extended-release IM injection formulation

Exclude: Short-acting injectables, as they are generally only used in emergent, acute conditions and on a short term basis (hours to days)

  - Key Question 2: Psychosocial and other non-pharmacological interventions (patients may also be on pharmacologic interventions):
    - Assertive community treatment (ACT)
    - Assisted outpatient treatment
    - Cognitive behavioral therapy for psychosis (CBTp)
    - Cognitive remediation
    - Family psychoeducation and behavior management
    - Illness management and recovery (IMR) toolkit
    - Integrated treatment for co-occurring (dual diagnosis) substance use disorders
    - Intensive case management
    - Peer support and peer-delivered services

- Social skills training
- Supported education
- Supported employment
- Supported housing
- Other interventions, including combinations

Note: Limited to the most commonly-used interventions today relevant to U.S. practices.

- Comparators:
  - Key Question 1:
    - Head to head comparisons: three first-generation antipsychotics (listed in table above) and all FDA-approved second-generation antipsychotics
    - Exclude: 1st vs. 1st generation drug comparisons
  - Key Question 2:
    - Antipsychotic drugs (alone)
    - Usual care/standard care/treatment as usual/waitlist, as defined in the studies, or nonactive/sham comparators, which act as a control for the non-specific effects of therapy
      - It is anticipated that both groups (treatment and usual care) may be receiving drug treatment as the baseline treatment.

Note: Comparisons for Key Question 2 are limited to usual care or drug interventions in order to evaluate which interventions are effective in general. Head to head comparisons are excluded for Key Question 2 because the intention here is not to demonstrate which interventions are better than others.

- Outcomes for each question:
  - Key Questions 1 and 2**:**
    *Benefits Outcomes:*
    - Functional
      - Improvements in social and occupational functioning
      - Enhanced level of independent or stable living situation
      - Reductions in legal system encounters
    - Reductions in self-harm, including suicide and suicide attempts
    - Improvements in health-related quality of life (using validated scales and including mental and physical health)
    - Ability to maintain treatment (e.g., treatment discontinuation/switching rate, time to discontinuation)
    - Improvements in core illness symptoms (e.g., delusions, hallucinations, disorganized thinking)
      - Rates of response, remission and relapse; speed and duration of response

        - Total scale scores of positive (i.e., delusions and hallucinations) and negative (i.e., passive or apathetic social withdrawal and blunted affect) symptoms
        - Reductions in agitation symptoms or aggressive behaviors
    - Changes in the status of co-occurring substance use disorder

Exclusions:

1. Re-hospitalization was not included because it is considered a flawed outcome measure for two reasons, 1) there is important variation in the indications for and length of psychiatric hospitalizations across time and in different settings, and 2) there is important variation across studies in how it is measured/evaluated, which may confound interpretation.
2. Neurocognitive testing is an intermediate outcome, rather than a patient-centered health outcome, and is excluded in favor of improvements in functioning that reflect cognition.

*Harms Outcomes:*

*KQ 1:*

- Overall adverse events (frequency of any adverse events reported in trials)
- Withdrawals due to adverse events, time to withdrawal due to adverse events
- Mortality (all-cause and cause-specific as defined by studies)
- Significant (major) adverse events
    - E.g., life threatening, result in long-term morbidity, or require medical intervention to treat, such as cerebro- or cardiovascular disease and related events, development of diabetes mellitus, diabetic ketoacidosis, neuroleptic malignant syndrome, seizures, extrapyramidal symptoms (EPS), tardive dyskinesia, clinically important weight changes, dyslipidemia, incidence and severity of sexual dysfunction, galactorrhea, amenorrhea, orthostatic hypotension and agranulocytosis

*KQ 2:*

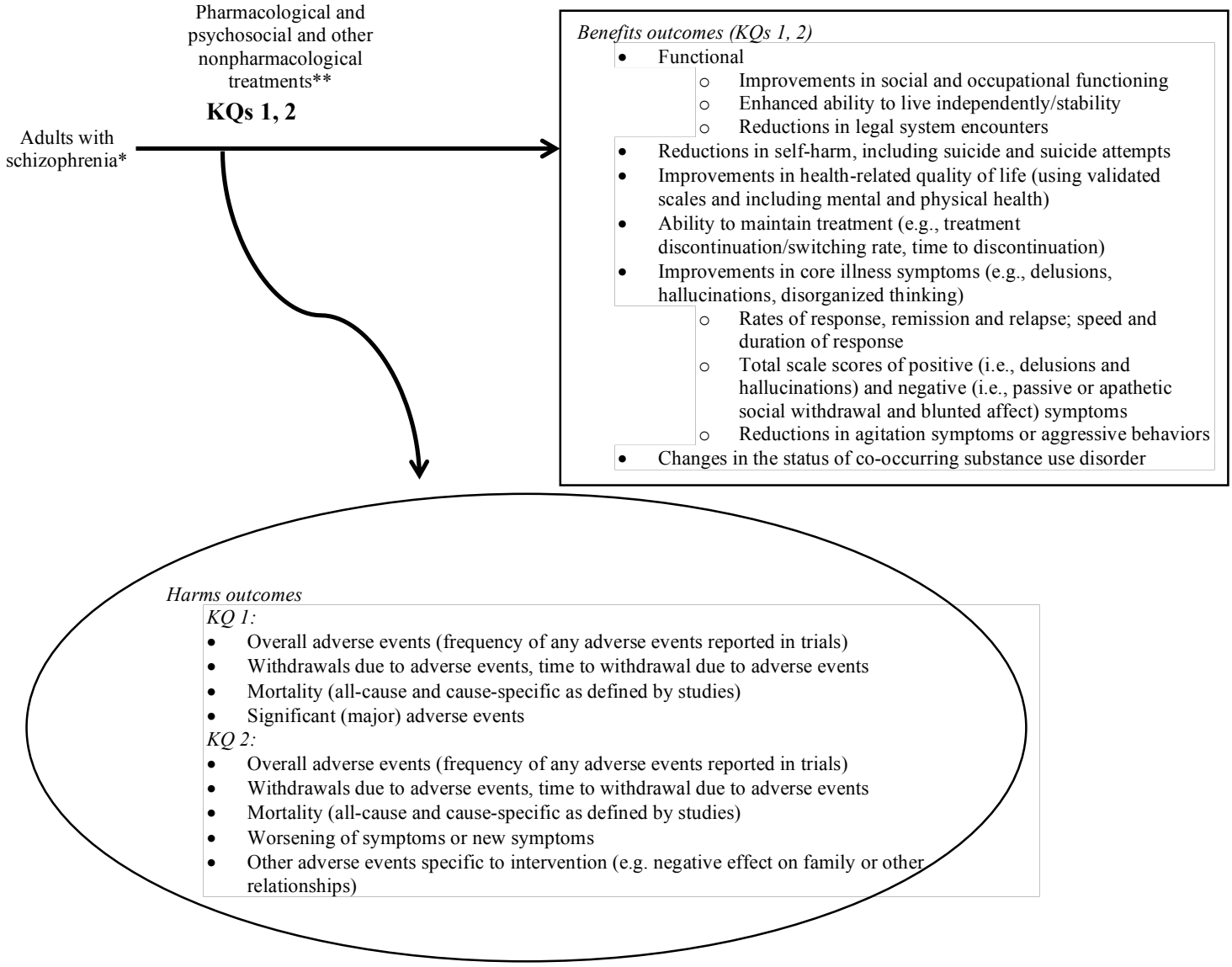
- Overall adverse events (frequency of any adverse events reported in trials)
- Withdrawals due to adverse events, time to withdrawal due to adverse events
- Mortality (all-cause and cause-specific as defined by studies)
- Outcomes reported as adverse events related to the intervention, e.g.:
    - New or worsening symptoms (e.g., anxiety or depression) using validated scales
    - Negative effect on family or other relationships

- Timing:
    - Minimum duration of follow-up: 12 weeks

- Study Designs
  - Key Questions 1 and 2:
    Best-evidence approach:
    - Recent, comprehensive, good- or fair-quality systematic reviews to be used as primary evidence, as well as RCTs published since systematic reviews
      - For benefits of any included intervention, systematic reviews of RCTs will be included
      - For harms of any included intervention, systematic reviews of observational studies to evaluate harms will be included in addition to reviews of trials.
    - If no systematic reviews available for particular interventions, RCTs will be included
    - Sample size of >50 for Key Question 2

- Settings:
  - U.S.-relevant, e.g. countries listed as "high" or "very high" on the United Nations International Human Development Index (HDI), and applicable to U.S. practices
  - Exclude: inpatient setting

## III. Analytic Framework




* Adults with a diagnosis of schizophrenia, including those with co-occurring substance use disorders, and including those experiencing a first episode of schizophrenia (including those with schizophreniform disorder).

- Pharmacological treatments:
  - At least 90 percent of patients must have been diagnosed with schizophrenia.
  - For studies specifically on harms of antipsychotic drugs, populations can be mixed-diagnoses, as the harms are not population-specific
- Psychosocial and other nonpharmacological treatments: 50 percent of patients must have been diagnosed with a schizophrenia spectrum disorder diagnosis (i.e., schizophrenia, schizoaffective disorder, or schizophreniform disorder)

**Pharmacological treatments include FDA-approved 2nd generation and selected 1st generation antipsychotics. Psychosocial and other nonpharmacological treatments include assertive community treatment (ACT), assisted outpatient treatment, cognitive behavioral therapy for psychosis (CBTp), cognitive remediation, family psychoeducation and behavior management, illness management and recovery (IMR) toolkit, integrated treatment for co-occurring (dual diagnosis) substance use disorders, intensive case management, peer support and peer-delivered services, social skills training, supported education, supported employment, supported housing, and other commonly-used interventions, including combinations.

## IV. Methods

### Criteria for Inclusion/Exclusion of Studies in the Review -

The criteria for inclusion and exclusion of studies will be based on the Key Question and are described in the previous PICOTS section.

Below are additional details on the scope of this project:

Study Designs: The decision to utilize a best-evidence approach and focus on existing systematic reviews was made due the combination of the high volume of trials expected, and the availability of high-quality, comprehensive evidence reviews for this topic identified during topic refinement. We will follow the AHRQ EPC guidance on using existing systematic reviews.[27] Specifically, systematic reviews will be used if they address a key question (and include all eligible interventions for the key question), include studies that meet the PICOTS as defined above, and are assessed as being at low risk of bias, according to the AMSTAR quality assessment tool.[28,29] If systematic reviews are included, we will update findings with any new primary trials identified in our searches. For key outcomes, we will update meta-analyses. For secondary outcomes, we will summarize systematic review findings and newer trial findings; if newer evidence contradicts older evidence we will investigate potential reasons. If multiple systematic reviews are relevant and low risk of bias, we will focus on the findings from the most recent reviews and evaluate areas of consistency and inconsistency across the reviews.[27] If systematic reviews for particular interventions or outcomes are not available, trials that satisfy the PICOT requirements will be included.

Non-English Language Studies: We will restrict to English-language articles, but will review English language abstracts of non-English language articles to identify studies that would otherwise meet inclusion criteria, in order to assess for the likelihood of language bias.

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

Publication Date Range:

For Key Question 1 on pharmacologic interventions, recent high quality systematic reviews directly addressing large portions of the Key Questions in the current review have been published and will be used as the starting point for the review.[30,31] Based on the search dates in these reviews, searches for trials and systematic reviews will begin in 2011 for 1st versus 2nd generation drugs and in 2013 for 2nd versus 2nd generation drugs. Starting the searches in January of 2011 and 2013, respectively, allows for multiple months of overlap of the new search dates with the search dates in the prior reviews.

For Key Question 2 on nonpharmacological interventions, search dates will not be restricted. Within these searches we will first identify the most recent, good-quality systematic reviews for particular interventions. Any trials identified in our searches that were published since the search dates in these reviews will also be included to update the included reviews.

Library searches will be updated while the draft report is posted for public comment and peer review to capture any new publications. Literature identified during the updated search, or through other methods described below, will be assessed by following the same process of dual review as all other studies considered for inclusion in the report. If any pertinent new literature is identified for inclusion in the report, it will be incorporated before the final submission of the report.

Literature Databases: Ovid MEDLINE, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, and PsycINFO will be searched to capture published literature.

Scientific Information Packets: The AHRQ Evidence-based Practice Center (EPC) Scientific Resource Center (SRC) will send email notification to relevant stakeholders about the opportunity to submit Scientific Information Packets (SIP) via the Effective Health Care (EHC) Web site for the pharmaceutical interventions listed in Key Question 1. These contain both published an unpublished evidence relevant to the review and will be reviewed according to the criteria and processes described for all evidence, below.

Hand Searching: Reference lists of included articles will also be reviewed for includable literature.

Grey literature: Searches for grey (unpublished) literature will include the SIPs submitted for pharmacological interventions and the ClinicalTrials.gov trial registry to identify trials of non-pharmacological interventions that have been completed.

Contacting Authors: In the event that information regarding methods or results appear to be omitted from the published results of a study, or if we are aware of unpublished data, we will query the authors to obtain this information.

Process for Selecting Studies: Pre-established criteria will be used to determine eligibility for inclusion and exclusion of abstracts in accordance with the AHRQ Methods Guide.[32] To ensure accuracy, all excluded abstracts will be dual reviewed. All citations deemed appropriate for inclusion by at least one of the reviewers will be retrieved. Each full-text article will be independently reviewed for eligibility by at least two team members, including any articles suggested by peer reviewers or that arise from the public posting or SIP processes. Any disagreements will be resolved by consensus.

Data Abstraction and Data Management -

After studies are selected for inclusion, the following data will be abstracted into pre-determined table templates: study design, year, setting, country, sample size, eligibility criteria, population and clinical characteristics, intervention characteristics, and results relevant to each key question as outlined in the previous PICOTS section, as well as other information. Information that will be abstracted that is relevant for assessing applicability will include the number of patients randomized relative to the number of patients enrolled, use of run-in or wash-out periods, and characteristics of the population, intervention, and care settings. All study data will be verified for accuracy and completeness by a second team member. A record of studies excluded at the full-text level with reasons for exclusion will be maintained.

Assessment of Methodological Risk of Bias of Individual Studies -

Predefined criteria will be used to assess the quality, which correlates with the risk of bias, of individual controlled trials, systematic reviews, and observational studies by using clearly defined templates and criteria as appropriate. Randomized controlled trials will be evaluated with appropriate criteria and methods developed by the Drug Effectiveness Review Project (DERP).[33] Systematic reviews will be assessed using the AMSTAR quality rating instrument.[29] These criteria and methods will be used in conjunction with the approach recommended in the chapter, Assessing the Risk of Bias of Individual Studies When Comparing Medical Interventions[28] in the AHRQ Methods Guide developed by the Agency for Healthcare Research and Quality. Studies will be rated as "good," "fair," or "poor," or as specified by the particular criteria.

Studies rated "good" quality will be considered to have the least risk of bias, and their results will be considered valid. Good-quality studies include clear descriptions of the population, setting, interventions, and comparison groups; a valid method for allocation of patients to treatment; low dropout rates and clear reporting of dropouts; appropriate means for preventing bias; and appropriate measurement of outcomes.

Studies rated "fair" quality will be susceptible to some bias, although not enough to invalidate the results. These studies may not meet all the criteria for a rating of good quality, but no flaw is likely to cause major bias. The study may be missing information, making it difficult to assess limitations and potential problems. The fair-quality category is broad, and studies with this rating will vary in their strengths and weaknesses. The results of some fair-quality studies are likely to be valid, while others may be only possibly valid.

Studies rated "poor" quality will have significant flaws that imply biases of various types that may invalidate the results. They will have a serious or "fatal" flaw in design, analysis, or reporting; large amounts of missing information; discrepancies in reporting; or serious problems in the delivery of the intervention. The results of these studies will be least as likely to reflect flaws in the study design as the true difference between the compared interventions. We will not exclude studies rated as being poor

in quality a priori, but poor-quality studies will be considered to be less reliable than higher quality studies when synthesizing the evidence, particularly if discrepancies between studies are present.

Each study evaluated will be dual-reviewed for quality by two team members. Any disagreements will be resolved by consensus. Included systematic reviews may use other methods to evaluate internal validity, such as the Cochrane risk of bias tool. In such cases, we will use the original rating system in the detailed assessments of the key questions.

### Assessing Key Question 2 Interventions for Evidence Sufficiency

Prior to synthesizing evidence for intervention categories in Key Question 2, we will determine bodies of evidence that are sufficient to draw conclusions per the procedure described below. As there are large, high quality systematic reviews of the interventions in Key Question 1 that have found the evidence sufficient on key outcomes in the past, this evidence will not undergo this sufficiency assessment process and will undergo the procedures described in the following sections.

> *Assessment procedure to ascertain sufficiency of evidence:* All studies will undergo quality assessment as described above. If the studies in a given intervention category, when grouped by similar PICO and study design, is insufficient based on the overall study limitations (derived from the quality of individual studies), we will not evaluate the other domains. For interventions with sufficient evidence based on quality of included studies, we will then assess precision of selected outcomes based on the optimal information size for selected outcomes. The outcomes will be prioritized based on input from TEP (listed on page 14) and the availability of outcomes measured in included studies. If the body of evidence for each intervention and outcome group based on insufficient quality and or precision, no further assessment will be done. For other bodies which have sufficient quality and precision, the rest of the strength of evidence assessment would follow the procedure described on page 15.

For studies included in intervention categories where the body of evidence is deemed insufficient, we will present tables of the main study characteristics, descriptive summary of findings and whether significant, study quality and funding source. These bodies of evidence will not undergo further synthesis, but will be summarized in the text, including reasons for being deemed insufficient

Those Key Question 2 intervention areas deemed sufficient on the selected key outcome will undergo full assessment and synthesis as described in the following sections.

Data Synthesis -

We will construct evidence tables identifying the study characteristics (as discussed above), results of interest, and quality ratings for all included studies, and summary tables to highlight the main findings. Good- and fair-quality systematic reviews and trials will be the focus of the results for each Key Question. Results from systematic reviews will be presented, followed by presentation of individual trials.

Qualitative data will be summarized in summary tables and as ranges and descriptive analysis, and interpretation of the results will be provided.

Meta-analyses will be conducted to summarize data and obtain more precise estimates on outcomes for which studies are homogeneous enough to provide a meaningful combined estimate. The feasibility of a quantitative synthesis will depend on the number and completeness of reported outcomes and a lack of heterogeneity among the reported results. To determine whether meta-analysis could be meaningfully performed, we will consider the quality of the studies and the heterogeneity among studies in design, patient population, interventions, and outcomes, and may conduct sensitivity analyses. The Key Questions are designed to assess the comparative effectiveness and harms by patient and treatment characteristics. We will examine the possibility for conducting network meta-analyses to provide estimates of comparative effect across the interventions (within a key question). We will compile and summarize study characteristics and investigate whether there are important differences in the distribution in characteristics that might modify the treatment effects. The evidence base and the geometry of the treatment network will be presented graphically. Network meta-analyses will be conducted using a Bayesian hierarchical model[34-36] and in all models we will control for variation in study duration and dose levels. The appropriateness of combining direct and indirect evidence and the consistency of the network will be assessed by checking specific loops and comparing consistency and inconsistency models overall. Inconsistency will be explored if detected. Treatment ranking will be obtained from the Bayesian models if the network is consistent. Given that such analyses are exploratory, they will be limited to key effectiveness outcomes. Sensitivity analyses will be conducted to explore heterogeneity. Examples include age (younger and older patients), first-episodes, and duration of disease.

Prioritization of outcomes to undergo strength of evidence ratings was determined with input from the Technical Expert Panel (TEP). The prioritized outcomes are listed below, per intervention area:

**Pharmacologic Interventions**

1. Functional outcomes (e.g., social, occupational)
2. Health-related quality of life (including physical)
3. Rates of response and/or remission
4. Mortality (all-cause and/or specific)
5. Reductions in self-harm, suicide and suicide attempts
6. Overall/any adverse events (rate or proportion)

7. Improvements in core illness symptoms scale score changes
8. Withdrawal due to adverse events

**Psychosocial and Other Nonpharmacologic Interventions**

1. Functional (e.g., social, occupational)
2. Health-related quality of life
3. Reductions in self-harm, suicide and suicide attempts
4. Rates of response and/or remission
5. Improvements in core illness symptoms scale score changes
6. Ability to maintain treatment
7. Rates of relapse
8. Outcomes reported as adverse events related to the intervention

The TEP will be consulted during the course of the review if there are additional clinical or methodological issues that arise which the review team thinks warrants discussion or advice.

Results will be presented as structured by the Key Questions, and any prioritized outcomes will be presented first.

## Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes

The strength of evidence for each key outcome will be initially assessed by one researcher for each key clinical outcome (see PICOTS) by using the approach described in the AHRQ EPC Methods Guide.[37] Key outcomes will be clinical (i.e., health outcomes) and will be selected based on input from the TEP. To ensure consistency and validity of the evaluation, the grades will be reviewed by a senior reviewer and any disagreements resolved through consensus:

- Study limitations (low, medium, or high level of study limitations based on the quality/risk of bias of individual studies)
- Consistency (consistent, inconsistent, or unknown/not applicable)
- Directness (direct or indirect)
- Precision (precise or imprecise)
- Reporting bias (suspected or undetected)

The strength of evidence will be assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale by evaluating and weighing the combined results of the above domains:

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

deficiencies. We believe that the findings are likely to be stable, but some doubt remains.

- Low — We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient — We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Assessing Applicability –

Applicability will be considered according to the approach described in the AHRQ EPC Methods Guide.[38] We will use the PICOTs framework to consider the applicability of the evidence base for each key question, for example examining the characteristics of the patient populations (e.g., age, severity and duration of illness); fidelity to nonpharmalogical interventions (e.g., modified interventions); and study setting (e.g. U.S. versus non-U.S. and clinical setting). Variability in the studies may limit the ability to generalize the results to other populations and setting.

## V. References


1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th Ed.). Arlington, VA: American Psychiatric Publishing; 2013.
2. Tandon R, Gaebel W, Barch DM, et al. Definition and description of schizophrenia in the DSM-5. Schizophr Res. 2013 Oct;150(1):3-10. doi: 10.1016/j.schres.2013.05.028. PMID: 23800613.
3. McGrath J, Saha S, Chant D, et al. Schizophrenia: a concise overview of incidence, prevalence, and mortality. Epidemiol Rev. 2008;30:67-76. PMID: 18480098.
4. American Psychiatric Association. Work Group on Schizophrenia; Lehman AF, Lieberman JA, Dixon LB, et al. Practice Guideline for the treatment of patients with schizophrenia, Second Edition. Am J Psychiatry. 2004;161:1-56. PMID: 15000267.
5. American Psychiatric Association. Dixon L, Perkins D, Calmes C. Guideline Watch: Practice Guideline for the treatment of patients with schizophrenia. Arlington, VA: American Psychiatric Association; September 2009. http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/schizophrenia-watch.pdf. Accessed December 6, 2015.
6. Dixon LB, Dickerson F, Bellack AS, et al. The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. Schizophr Bull. 2010 Jan;36(1):48-70. doi: 10.1093/schbul/sbp115. PMID: 19955389.
7. Ben-Zeev D, Davis KE, Kaiser S, et al. Mobile technologies among people with serious mental illness: opportunities for future services. Adm Policy Ment Health. 2013 Jul;40(4):340-3. doi: 10.1007/s10488-012-0424-x. PMID: 22648635.
8. Buchy L, Cadenhead KS, Cannon TD, et al. Substance use in individuals at clinical high risk of psychosis. Psychol Med. 2015;45(11):2275-84. PMID: 25727300.
9. Goerke D, Kumra S. Substance abuse and psychosis. Psychiatr Clin North Am. 2013;22(4):643-54. PMID: 24012078.
10. Mesholam-Gately RI, Gibson LE, Seidman LJ, et al. Schizophrenia and co-occurring substance use disorder: Reward, olfaction and clozapine. Schizophr Res. 2014;155(1-3):45-51. PMID: 24685823.
11. De Witte NA, Crunelle CL, Sabbe B, et al. Treatment for outpatients with comorbid schizophrenia and substance use disorders: a review. Eur Addict Res. 2014;20(3):105-14. doi: 10.1159/000355267. PMID: 24192558.
12. American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, et al. Consensus development conference on antipsychotic drugs and obesity and diabetes. Diabetes Care. 2004 Feb;27(2):596-601. PMID: 14747245.
13. Newcomer JW, Hennekens CH. Severe mental illness and risk of cardiovascular disease. JAMA. 2007;298(15):1794-6. PMID: 17940236.
14. Henderson D, Vincenzi B, Andrea NV, et al. Pathophysiological mechanisms of increased cardiometabolic risk in people with schizophrenia and other severe mental illnesses. Lancet Psychiatry. 2015;2(5):452-64. PMID: 26360288.
15. Marder SR, Essock SM, Miller AL, et al. Physical health monitoring of patients with schizophrenia. Am J Psychiatry. 2004;161(8):1334-49. PMID: 15285957.
16. McDonagh MS, Peterson K, Carson S, et al. Drug Class Review of Atypical Antipsychotic Drugs; Update 3. Portland, OR: Oregon Health & Science University Drug Effectiveness Review Project; 2010.
17. Mayo Clinic. Diseases and Conditions: Schizophrenia. http://www.mayoclinic.org/diseases-conditions/schizophrenia/basics/definition/CON-20021077. Accessed December 6, 2015.

18. Centers for Disease Control. Burden of Mental Illness. 2013. http://www.cdc.gov/mentalhealth/basics/burden.htm. Accessed December 6, 2015.

19. National Institute of Mental Health. Schizophrenia. http://www.nimh.nih.gov/health/topics/schizophrenia/index.shtml. Accessed December 6, 2015.

20. Gareri P, De Fazio P, Manfredi VG, et al. Use and safety of antipsychotics in behavioral disorders in elderly people with dementia. J Clin Psychopharmacol. 2014;34(1):109-23. PMID: 24158020.

21. Kales HC, Kim HM, Zivin K, et al. Risk of mortality among individual antipsychotics in patients with dementia. Am J Psychiatry. 2012;169(1):71-9. PMID: 22193526.

22. Schneider LS, Dagerman KS, Insel P. Risk of death with atypical antipsychotic drug treatment for dementia: meta-analysis of randomized placebo-controlled trials. JAMA. 2005 Oct 19;294(15):1934-43. doi: 10.1001/jama.294.15.1934. PMID: 16234500.

23. Gill SS, Bronskill SE, Normand SL, et al. Antipsychotic drug use and mortality in older adults with dementia. Ann Intern Med. 2007 Jun 5;146(11):775-86. PMID: 17548409.

24. Alexopoulos GS, Jeste DV, Chung H, et al. The expert consensus guideline series. Treatment of dementia and its behavioral disturbances. Introduction: methods, commentary, and summary. Postgrad Med. 2005 Jan;Spec No:6-22. PMID: 17203561.

25. Jeste DV, Rockwell E, Harris MJ, et al. Conventional vs. newer antipsychotics in elderly patients. Am J Geriatr Psychiatry. 1999 Winter;7(1):70-6. PMID: 9919323.

26. Kreyenbuhl J, Buchanan RW, Dickerson FB, et al. The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2009. Schizophrenia Bulletin. 2010 Jan;36(1):94-103. doi: 10.1093/schbul/sbp130. PMID: 19955388.

27. White C, Ip S, McPheeters M, et al. Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. Rockville, MD; 2009. http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318. Accessed July 1, 2016.

28. Agency for Healthcare Research and Quality. Methods guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(13)-EHC063-EF. Rockville MD: 2013.

29. Shea B, Hamel C, Wells G, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epideomiol. 2009;62(10):1013-20. PMID: 19230606.

30. Abou-Setta AM, Mousavi SS, Spooner C, et al. First-generation versus second-generation antipsychotics in adults: comparative effectiveness. Comparative Effectiveness Review No. 63. Rockville, MD: Agency for Healthcare Research and Quality; 2012.

31. McDonagh M, Peterson K, Fu R, et al. Drug Class Review: Second Generation Antipsychotic Drugs. Final Report - Update 4. Prepared by the Pacific Northwest Evidence–based Practice Center for the Drug Effectiveness Review Project. Portland, OR: Oregon Health & Science University; 2013.

32. McDonagh M, Peterson K, Raina P, et al. Avoiding Bias in Selecting Studies. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Oregon Health & Science University, McMaster University, and Southern California Evidence-based Practice Centers under Contract No. 290-2007-10057-I.) AHRQ Publication No. 13-EHC045-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm Accessed July 1, 2016.

33. McDonagh M, Jonas D, Gartlehner G, et al. Methods for the Drug Effectiveness Review Project. BMC Med Res Methodol. 2012;12(1):140.

34. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. J Amer Statist Assoc. 2006;101:447-59.

35. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004;23:3105-24. DOI:10.1002/sim.875.

36. Salanti G, Higgins JP, Ades AE, et al. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008;17:279-301. DOI:10.1177/0962280207080643.

37. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of EvidenceWhen Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I). AHRQ Publication No. 13(14)-EHC130-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm Accessed July 1, 2016.

38. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1198-207. doi: 10.1016/j.jclinepi.2010.11.021. PMID: 21463926.

## VI. Definition of Terms

Not applicable.

## VII. Summary of Protocol Amendments

| Date | Section | Original Protocol | Revised Protocol | Rationale |
|---|---|---|---|---|
| September 21, 2016 | Page 13 | Previously, all studies for all intervention areas in Key Question 2 were to undergo full data abstraction and synthesis procedures. | The section on "Assessing Key Question 2 Interventions for Evidence Sufficiency" was added, which adds a preliminary step of assessing the intervention area for evidence sufficiency based on a systematic procedure. Those areas deemed insufficient would undergo limited data abstraction and synthesis procedures. (See page 13 for details.) | This change was made due to a very large evidence base, limited timeline and resources, and limited evidence for the intervention areas in Key Question 2. The intent was to be able to include interventions with limited evidence, and provide adequate information to result in a meaningful report. |

## VIII. Review of Key Questions

AHRQ posted the key questions on the Effective Health Care Website for public comment. The EPC refined and finalized the key questions after review of the public comments, and input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the key questions are specific and relevant.

## IX. Key Informants

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

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

## X. Technical Experts

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

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

## XI. Peer Reviewers

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

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

## XII. EPC Team Disclosures

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

## XIII. Role of the Funder

This project was funded under Contract No. HHSA290201500009I, Task Order No. 5 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.