Evidence-based Practice Center Systematic Review Protocol

**Project Title: Strategies to Improve Mental Health Care for Children and Adolescents**

Initial Publication Date if applicable: December 30, 2014

Amendment Date(s) if applicable: December 11, 2015

(Amendments Details–see Section VII)

## I. Background and Objectives for the Systematic Review

Approximately one in five children and adolescents living in the United States has one or more mental, emotional, or behavioral health disorders according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) criteria in any given year.[1] These disorders contribute to problems with family, peers, and academic functioning; comorbidity (including other mental and substance use disorders and chronic health conditions); and reduced quality of life; they also increase the risk of involvement with the criminal justice system and other risk-taking behaviors and suicide.[2]

The evidence base for pediatric mental health interventions that target mood disorders, anxiety disorders, disruptive behavior disorders, psychotic disorders, eating disorders, and substance use disorders continues to grow.[3,4] Despite advances in the evidence base, outcomes for children with mental health problems remain suboptimal because of issues with access to care, failure of systems and providers to adopt interventions with proven efficacy (e.g., evidence-based practices [EBPs]), and variability in the quality of mental health care received. Studies using nationally representative data on U.S. adolescents show that only approximately one in five children with mental health problems receives services, and only one-third of treatment episodes are considered minimally adequate (at least four visits with psychotropic medication or at least eight visits without psychotropic medication).[5-7] The current health care system continues to provide fragmented care to children in numerous uncoordinated systems, rendering inefficient delivery of needed services.[8] Other issues include providers not having the time available or knowledge/training to identify mental health problems and treat or refer accordingly.[9]

Several key publications in the mid- to late-1990s suggested that usual care in children's mental health had, at best, null[10] and sometimes harmful effects.[11] With the proliferation of EBPs addressing childhood mental disorders,[3,12] strategies to implement or disseminate these interventions targeted a change in the organization and delivery of mental health services.[13,14] These strategies sought to improve the quality of care by closing the gap between research evidence and practice.[15,16]

This review will focus on strategies that aim to improve the quality of mental health care rather than to determine the efficacy of interventions or quality of specific EBPs. We plan to focus on dissemination, implementation, and quality improvement (D/I/QI) strategies targeting providers, organizations, or systems that care for children and adolescents with

mental health problems. Various studies have adopted the framework presented in the Institute of Medicine's (IOM's) landmark *Crossing the Quality Chasm* report[16] to define a quality improvement (QI) strategy. For this review, we plan to define QI strategies as those that target providers (e.g., via education, training, and supervision) and organizations (e.g., via financial incentives, regulation, and policies) that provide mental health care to children and adolescents, with the ultimate goal of improving the quality of care.[17,18]

Some consider dissemination/implementation (D/I) strategies as a particular subset of QI initiatives, although the field is so new that terminology has not yet been standardized.[19] Dissemination strategies can be differentiated from implementation strategies using the definition constructed by McCormack and colleagues:[20] dissemination is the active and targeted distribution of information and interventions to a specific public health or clinical practice audience via determined channels using planned strategies with the intent to spread knowledge and associated evidence-based interventions to enhance the adoption and the implementation of the information and/or intervention; implementation is the use of strategies to integrate evidence-based health interventions (e.g., EBPs) and change practice patterns within specific settings. Although the ultimate goal of D/I/QI strategies is to improve patient-related outcomes for children and adolescents with mental health problems, intermediate outcomes of D/I/QI strategies include changes to health care systems, organizations, and providers that provide mental health care. Recent D/I theoretical frameworks posit that effective and sustainable implementation operates through multiple nested levels.[21,22] These levels typically include the macro-environment (i.e., state), organization or system (i.e., specialty mental health clinic), program (i.e., selected intervention), provider (i.e., clinicians), and patients (children and families). Outcomes across these various levels are interrelated. For instance, changes in intermediate outcomes such as provider attitudes[23] or organizational climate[24] may influence the successful adoption of and fidelity to EBPs, which in turn influence patient outcomes, such as behavior or quality of life.

### ***Potential Moderators of Strategy Effectiveness***

Several frameworks have been developed to study how certain variables, including contextual factors, influence the effectiveness of the D/I or QI strategy.[25-29] One framework commonly used to study implementation research, the Consolidated Framework for Implementation Research (CFIR),[27] comprises five major domains: intervention characteristics (e.g., evidence strength and quality), inner setting (e.g., culture, leadership, engagement), outer setting (e.g., patient needs and resources, external policies and incentives), the characteristics of involved individuals (provider training, experience), and the process by which implementation is accomplished (e.g., plan, evaluate, and reflect). We will use the CFIR as an organizing framework for moderators of strategy effectiveness. In addition to the five domains, we have added another category, characteristics of the patient. Research is under way to determine the most important contextual factors for effective implementation of mental health strategies for children and adolescents.[30]

One challenge related to the generalizability and applicability of D/I/QI research is the diversity of outpatient settings (inner setting) for children's mental health services, which include primary care, schools, specialty mental health, emergency rooms, or,

increasingly, home-based service provision. Research suggests that the organizational factors of the clinical (inner) setting influence outcomes, and many have argued that these unique factors should be studied within the context of D/I/QI research.[31,32] The external context or outer setting also matters. As the health care environment continues to change, it has experienced fragmentation due to the increasing fiscal pressures that many states and local communities face. Thus, the diverse and changing clinical structure underlying settings where D/I/QI strategies are tested will make understanding the impact of context on intervention effectiveness important to study.

**Rationale for Evidence Review**

The recent proliferation of D/I/QI strategies for children and adolescents with mental health problems indicates that the existing body of evidence stands poised for an objective systematic review. Decisionmakers are in critical need of information about D/I/QI strategies to improve children's mental health care. An improved understanding of the comparative benefits, harms, and modifiers of the available strategies to improve mental health care for children and adolescents may help guide providers of care, administrators of care facilities, organizations, and health systems and inform insurance coverage decisions and other policy decisionmaking for children and adolescents with mental health care needs.

Two recent systematic reviews have been published on this topic. Barwick and colleagues published a review in 2012 that focused on knowledge translation interventions/strategies related to the delivery, organization, or receipt of child and youth mental health services.[33] Most focused on practitioner or teacher training for behavior change. This systematic review excluded studies of children with substance abuse. The second systematic review on this topic, conducted by Novins and colleagues[34] and published in 2013, focused on the dissemination and implementation of EBPs for children and adolescent mental health.[34] The scope of our proposed review will add to this evidence base by focusing more broadly on QI strategies, searching for harms, and exploring differential effectiveness of contextual factors and other characteristics. The newness of the field, coupled with the importance of moving effective interventions into usual care settings, underscores the critical nature of the proposed review to improving care for children and adolescents with mental and behavioral health problems.

We propose to focus our review on D/I/QI strategies that target providers and health care organizations that provide mental health services for children and adolescents experiencing already existing mental health symptoms (i.e., we would not include strategies such as the implementation of educational interventions for reading disorders). We also will limit our review of implementation strategies to those focusing on EBPs. For the purpose of defining EBP, we propose to rely on the minimum requirements set forth by the Substance Abuse and Mental Health Services Administration's (SAMHSA's) National Registry of Evidence-based Programs and Practices (NREPP) (www.nrepp.samhsa.gov). These criteria specify that the intervention needs to have produced one or more positive behavioral outcomes in at least one study using an experimental or quasi-experimental design with results published in a peer-reviewed journal or similar publication. In addition, implementation materials, training and support resources, and quality assurance procedures for these interventions need to be ready for use by the public.

## II. The Key Questions

As listed below, the review will address Key Questions (KQs) that were revised after posting for comment. Thus, these KQs reflect input received from the public. The comments generally focused on ensuring that D/I/QI strategies had distinct definitions, accounting for the importance of considering comorbid conditions experienced by the child or adolescent when examining strategy effectiveness, and identifying how the moderators affect outcomes of a particular strategy when answering KQ 3. Our revised KQs include:

Key Question 1: What is the effectiveness of dissemination, implementation, and quality improvement (D/I/QI) strategies employed in outpatient settings by health care providers, organizations, or systems that care for children and adolescents with mental health problems to improve (a) intermediate patient, provider, or system outcomes and (b) final health or patient-centered outcomes?

Key Question 2: What are the harms of these D/I/QI mental health strategies?

Key Question 3: Do characteristics of the child or adolescent or contextual factors (e.g., characteristics of providers, organizations, or systems; intervention characteristics, setting; or process) modify the effectiveness or harms of D/I/QI mental health strategies and, if so, how?

**Populations**: Health care systems, organizations, and providers that care for children and adolescents with mental health problems including behavioral disorders (e.g., conduct disorder, attention deficit hyperactivity disorder, oppositional defiant disorder) and substance use disorders

**Interventions: D/I/QI s**trategies targeting health care systems or providers to improve the quality of care for children and adolescents with mental health problems

1. Dissemination: strategies used to disseminate evidence through increasing the reach, people's motivation to use and apply EBPs (defined based on the minimum criteria set forth by SAMHSA's NREPP), and/or people's ability to use and apply EBPs. Examples include (1) strategies to increase the reach of the evidence (e.g., social media, interpersonal outreach); (2) strategies to increase people's motivation to use and apply the evidence (e.g., use of opinion leaders, champions, social networks); (3) strategies to increase people's ability to use and apply the evidence (e.g., additional resources, skills building); and (4) strategies that use a multipronged approach with any of the previously described dissemination strategies (e.g., social marketing, academic detailing).[20]
2. Implementation: strategies used to adopt and integrate EBPs (defined based on the minimum criteria set forth by SAMHSA's NREPP) into routine care (e.g., strategies to integrate evidence-based health interventions and change practice patterns). Examples of implementation strategies that vary by method of implementation facilitation include (1) plan strategies, (2) educate strategies, (3) finance strategies, (4) restructure strategies, (5) quality management strategies, and (6) attend to policy context.[35]
3. Quality improvement: strategies targeting systems and providers of mental health care to children and adolescents with the goal of improved quality of care. Examples of

quality improvement strategies include (1) provider targets: audit and feedback, facilitated relay of clinical data to providers, pay for performance, and provider reminder systems; (2) organization/system targets: changes to the organization including case management, changing from paper to computer systems, increased staffing, changes in reimbursement schemes.[16,36]

**Comparators:**

1. Usual care, as defined by studies
2. Other D/I/QI strategy

**Outcomes:**

Intermediate outcomes (at least one intermediate outcome required for KQs 1, 3)

1. Provider
   - satisfaction with or acceptability of approach
   - protocol adherence/program model fidelity
   - competence/skills
2. System
   - feasibility
   - uptake
   - timeliness
   - penetration
   - sustainability
   - costs
3. Patient
   - access to care
   - satisfaction
   - treatment engagement
   - therapeutic alliance with provider

Final health or patient-centered outcomes (at least one final health or patient-centered outcome is required for KQs 1, 3 unless strategy uses an intervention that is an EBP)

1. mental health symptoms, syndromes, or disorders
2. comorbidity
3. mortality
4. socialization skills and behavior
5. functional status
6. quality of life
7. service utilization (e.g., visits, hospitalizations)

Harms of intervention(s) (KQs 2, 3)

1. Patient
   - lower treatment engagement/increased dropouts
   - negative impact on therapeutic relationship

- side effects of intervention (e.g., adverse events, weight changes, suicidality)
- patient dissatisfaction with care

2. Provider
   - burnout/exhaustion
   - turnover
   - resistance to D/I/QI intervention
3. Organization
   - cost
   - failure to sustain EBP
   - resistance to change

**Timing:**

Any duration of followup

**Setting:**

1. Studies conducted in highly developed countries ("very high" human development index per the United Nations Development Programme)
2. Outpatient settings serving children and adolescents with mental health problems (primary care, specialty care, emergency rooms, community mental health centers, integrated care settings, federally qualified health centers, school-based mental health care, home-based care)

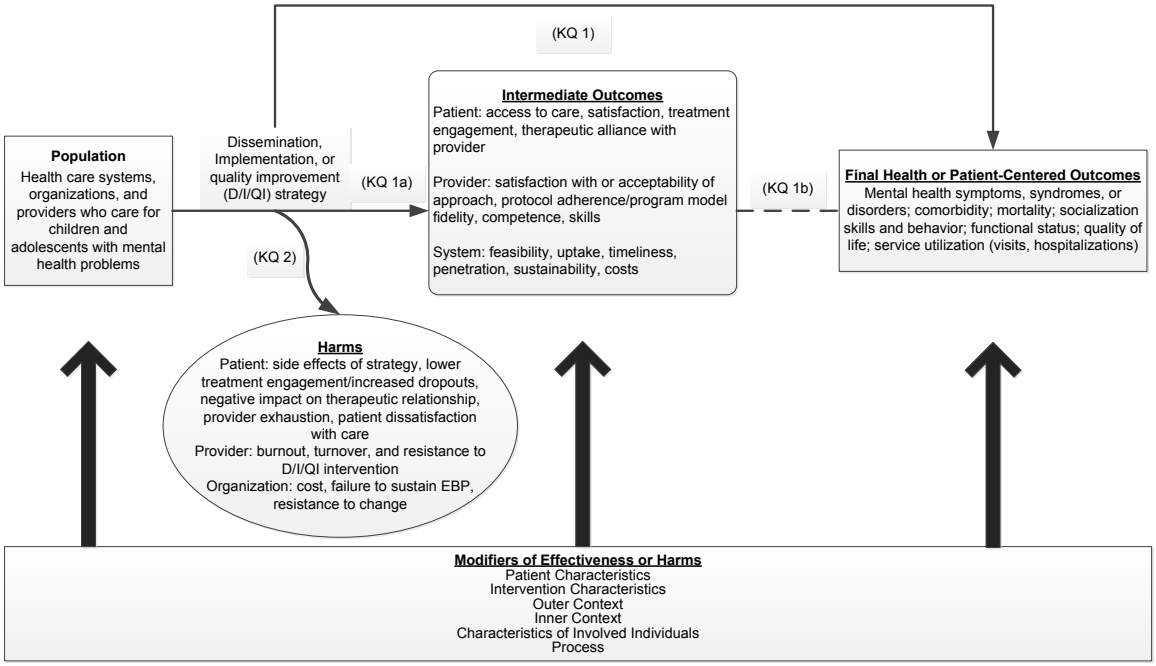
**Moderators:**

1. Patient characteristics (age, gender, cognitive functioning, diagnosis/severity of mental health problem, comorbid conditions, cotreatments, race/ethnicity)
2. Intervention characteristics (complexity, manualized or not, intensity/frequency/duration, adjustment of intervention to fit context)
3. Outer setting (external policy, incentives, availability of alternative care systems)
4. Inner setting/organizational factors (type of outpatient setting, structure/size, culture, implementation climate, readiness of organization for implementation)
5. Characteristics of involved individuals (provider type, knowledge, beliefs, self-efficacy, leadership, education, certifications, accreditation policies, standards, and years of practice)
6. Process characteristics (fidelity to the planned strategy, fidelity to the EBP, use of champions or supervision/oversight)
7. Other (length of followup)

## III. Analytic Framework

The relationship between the patient population, interventions, comparators, outcomes, and timing of outcomes assessment (PICOT) is depicted in relation to the KQs (Figure 1).

**Figure 1. Analytic framework for strategies to improve mental health care in children and adolescents**




## IV. Methods

**Criteria for Inclusion/Exclusion of Studies in the Review:** We specified our inclusion and exclusion criteria based on the populations, interventions, comparators (control intervention), outcomes, timing, and settings identified through the topic refinement exercise (Table 1). Of note, the population includes health care systems, organizations, and providers that care for children and adolescents or mixed populations with mental health problems. Our exclusion of non–English-language studies is based on limitations of time and resources. However, we will examine English language abstracts of non–English-language studies to assess the potential size of the literature that would be missed through this approach. We will exclude study designs without control (or comparison) groups to ensure that our pool of included studies can inform the causal link between the strategy and outcomes for all KQs. For KQ 1 studies of benefits and KQ 3 studies of moderators of benefits, we will include only randomized controlled trials (RCTs) (standard, clustered, stepped-wedge), controlled clinical trials (CCTs, not randomized), systematic reviews, or meta-analyses. If we find no evidence to answer our KQs using these study designs, we will consider other designs, specifically, cohort studies (prospective, retrospective, and historical control), interrupted time-series, and case-control studies that meet all other inclusion and exclusion criteria. For KQ 2 and KQ 3 studies on moderators of harms, we will include experimental studies noted above, interrupted time-series, and observational evidence from prospective cohort studies, retrospective cohort studies, and case-control studies that meet all other inclusion and exclusion criteria.

**Table 1. Inclusion/exclusion criteria**

| | Criteria | |
|---|---|---|
| **Category** | **Inclusion** | **Exclusion** |
| Population | Health care systems, organizations, and providers that care for children and adolescents or mixed populations with mental health problems | Health care systems, organizations, and providers that care for adults 18 years of age or older only<br>Health care systems, organizations, and providers that care for children and adolescents with developmental disorders only |
| Geography | Countries with a very high human development index (HDI) | Countries with high, medium, low, or very low HDI |
| Date of search | All | None |
| Study duration | No limit | None |
| Settings | Outpatient settings serving children and adolescents with mental health problems (primary care, specialty care, emergency rooms, community mental health centers, integrated care settings, federally qualified health centers, school-based mental health care, home-based care) | Inpatient or residential treatment settings, drug treatment programs, settings in jail/prisons |
| Interventions (Strategies) | • Dissemination strategies (e.g., strategies to enhance the adoption and the implementation of evidence-based interventions that meet NREPP inclusion criteria)<br>• Implementation strategies(e.g., strategies to integrate evidence-based interventions that meet NREPP inclusion criteria with the goal of changing practice patterns)<br>• QI strategies (e.g., strategies targeting systems and providers of mental health care to children and adolescents with the goal of improved quality of care) | Interventions targeting patients only, drug interventions only, interventions not otherwise described in inclusion criteria |
| Comparator | Any control strategy, including usual care or different variants of the same intervention | None |

**Table 1. Inclusion/exclusion criteria (continued)**

| | Criteria | |
|---|---|---|
| **Category** | **Inclusion** | **Exclusion** |
| Outcomes | Intermediate outcomes (at least one intermediate outcome is required for KQs 1, 3)<br>• Provider<br>• satisfaction with or acceptability of approach<br>• protocol adherence/program model fidelity<br>• competence/skills<br>• System<br>• feasibility<br>• uptake<br>• timeliness<br>• penetration<br>• sustainability<br>• resources (including costs)<br>• Patient<br>• access to care<br>• satisfaction<br>• treatment engagement<br>• therapeutic alliance with provider<br>Final health or patient-centered outcomes (at least one final health or patient-centered outcome is required for KQs 1, 3 unless strategy uses an intervention that is an EBP)<br>• Change in mental health status, including symptom change, response, remission, relapse, and recurrence<br>• Comorbid physical health conditions, substance use problems, developmental disorders, other mental health problems<br>• Mortality<br>• Socialization skills and behavior<br>• Functional status<br>• Quality of life<br>• Service utilization (e.g., visits, hospitalizations) | All outcomes not otherwise specified |

**Table 1. Inclusion/exclusion criteria (continued)**

| Category | Criteria | |
|---|---|---|
| | **Inclusion** | **Exclusion** |
| Outcomes (continued) | Harms of strategy<br>1. Patient<br>• lower treatment engagement/increased dropouts<br>• negative impact on therapeutic relationship<br>• side effects of evidence-based practice incorporated into strategy (e.g., adverse events, suicidality)<br>• patient dissatisfaction with care<br>2. Provider<br>• burnout/exhaustion<br>• turnover<br>• resistance to D/I/QI intervention<br>3. Organization<br>• cost<br>• failure to sustain EBP<br>• resistance to change | |
| Timing of outcome measurement | All | None |
| Publication language | English | All other languages |
| Study design | Original research with eligible study designs to include:<br>KQs 1, 3 (benefits) [a]:<br>• RCTs<br>• CCTs<br>• Systematic review and meta-analyses<br>KQs 2, 3 (harms):<br>• RCTs<br>• CCTs<br>• Systematic review and meta-analyses<br>• Cohort studies<br>• Interrupted time series<br>• Case-control studies | • Case series<br>• Case reports<br>• Nonsystematic reviews<br>• Cross-sectional studies<br>• Before and after studies without time series data<br>• Other designs without a control or comparison group |

**Table 1. Inclusion/exclusion criteria (continued)**

| | Criteria | |
|---|---|---|
| **Category** | **Inclusion** | **Exclusion** |
| Publication type | Any publication reporting primary data | Publications not reporting primary data |

CCT: controlled clinical trial; EBP = evidence-based practice; D/I/QI = dissemination, implementation, and quality improvement; HDI = human development index; KQ = Key Question; NREPP = National Registry of Evidence-based Programs and Practices; QI = quality improvement; RCT = randomized controlled trial.

[a] Will include cohort, interrupted time-series, and case-control studies only if no RCT/CCT/systematic reviews/meta-analysis evidence is found.

Searching for the Evidence: Literature Search Strategies for Identifying Relevant Studies to Answer the KQs: We will systematically search, review, and analyze the scientific evidence for each KQ. We will take the following steps to perform the literature search. To identify articles relevant to each KQ, we will begin with a focused MEDLINE® search for eligible interventions using a combination of medical subject headings (MeSH®) and title and abstract keywords, limiting the search to human-only studies (preliminary search string for MEDLINE search is presented in the Appendix) . We will also search the Cochrane Library, PsycINFO, and CINAHL using analogous search terms. These searches will include RCTs, CCTs, and systematic reviews/meta-analyses. We selected these databases based on preliminary searches and consultation with content experts. We will conduct quality checks to ensure that the search identifies known studies (i.e., studies identified during topic nomination and refinement). If we do not identify the known studies, we will revise and rerun our searches.

In addition, we will search the "gray literature" for unpublished studies relevant to this review and will include studies that meet all the inclusion criteria and contain enough methodological information to assess risk of bias. Potential sources of gray literature include ClinicalTrials.gov, the World Health Organization's International Clinical Trials Registry Platform, Health Services Research Projects in Progress, the National Institutes of Health Research Portfolio Online Reporting Tools, the Database of Promoting Health Effectiveness Reviews, the New York Academy of Medicine Grey Literature Report, and CMS.gov. To avoid retrieval bias, we will manually search the reference lists of landmark studies and background articles on this topic to look for any relevant citations that our electronic searches might have missed.

We will also conduct an updated literature search (of the same databases searched initially) concurrent with the peer review process. We will investigate any literature the Technical Expert Panel (TEP), peer reviewers or the public suggest and, if appropriate, will incorporate additional studies into the final review. The appropriateness of those studies will be determined using the methods described above.

**Data Abstraction and Data Management:** Two trained research team members will independently review all titles and abstracts identified through searches for eligibility against our inclusion/exclusion criteria. Studies marked for possible inclusion by either reviewer will undergo a full-text review. For studies without adequate information to determine inclusion or exclusion, we will retrieve the full text and then make the

determination. We will track all results in an EndNote® bibliographic database (Thomson Reuters, New York, NY).

We will retrieve and review the full text of all titles included during the title/abstract review phase. Two trained team members will independently review each full-text article for inclusion or exclusion based on the eligibility criteria described above. If both reviewers agree that a study does not meet the eligibility criteria, we will exclude the study. If the reviewers disagree, conflicts will be resolved by discussion and consensus or by consulting a third member of the review team. As described above, all results will be tracked in an EndNote database. We will record the reason that each excluded full-text publication did not satisfy the eligibility criteria so that we can later compile a comprehensive list of such studies.

For studies that meet our inclusion criteria, we will abstract important information into evidence tables. We will design data abstraction forms to gather pertinent information from each article, including, strategies (including EBP interventions), characteristics of strategy target (systems, organizations, providers), comparators, settings, characteristics of the children/adolescents with mental health problems served by the system/organization/provider targeted, study designs, analysis methods, and results. The forms will be compatible with criteria for inclusion in the Agency for Healthcare Research and Quality's Systematic Review Data Repository. Trained reviewers will extract the relevant data from each included article into the evidence tables. A second member of the team will review all data abstractions for completeness and accuracy. We identify studies that address KQ 3, moderators of strategy effectiveness or harms, as those that compare effectiveness or harms of a particular strategy between two or more levels of the moderator (e.g., contextual or other variable of interest). For example, we will include studies that compare effectiveness between levels of a potential moderator via interaction analysis. For example, we would include studies that compare the effectiveness of a particular strategy used in a school-based setting with a primary care setting (where inner setting [type of outpatient setting] is the moderator).

For systematic reviews, we will use the five-step process described in the AHRQ Methods Guide[37] to assess the relevance and quality of the systematic review and to determine how to use the information provided. We will incorporate existing systematic reviews or use them to replace all or part of the de novo process only if they are fully relevant and of high quality. Reviews that do not meet these criteria can be used to cross-check references. Systematic reviews meeting relevance and quality criteria can be used to refine the search strategy and/or the summarized evidence can be incorporated into our review.

**Assessment of Methodological Risk of Bias of Individual Studies**: To assess the risk of bias (internal validity) of studies, we will use predefined design-specific criteria based on guidance in the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (Methods Guide).[38] We will evaluate the risk of bias of RCTs using the Cochrane Risk of Bias Tool. For non-RCTs, we will use a tool developed by Viswanathan et al.[39] Should an updated tool for observational studies from the Cochrane Collaboration (currently in development) become available, we will attempt to use it. For systematic reviews, we will use a modified AMSTAR (Assessment of Multiple Systematic Reviews) instrument[40] to assess the quality of each relevant review. Minimum eligibility criteria for

systematic reviews will include an explicit description of search strategy used and determination that the search strategy is adequate, application of pre-defined eligibility criteria and risk of bias assessment for all included studies, and synthesis of the results presented.

In general terms, results of a study with low risk of bias are considered to be valid. A study with medium risk of bias is susceptible to some bias but probably not sufficient to invalidate its results. A study with high risk of bias has significant methodological flaws (i.e., stemming from serious errors in design or analysis) that may invalidate its results. We will consider the risk of bias for each relevant outcome of a study.

Two independent reviewers will assess the risk of bias for each study. Disagreements between the two reviewers will be resolved by discussion and consensus or by consulting a third member of the team. We will rate studies that meet all criteria as having "low risk of bias." "Medium risk of bias" ratings will be given to studies where raters have some confidence that the results represent the true treatment effect; that is, although the study is susceptible to some bias, the problems are not considered sufficient to invalidate the results (i.e., no flaw is likely to cause major bias). We will give a "high risk of bias" rating to studies that have a fatal flaw (defined as a methodological shortcoming that leads to a very high risk of bias) in one or more categories.

**Data Synthesis:** If we find five or more similar studies that use a common design (all RCTs or all cohort) for a comparison of interest, we will consider quantitative analysis (i.e., meta-analysis) of the data from those studies.[41] We will also consider conducting mixed treatment comparisons meta-analysis using Bayesian methods to compare interventions with one another if we identify a sufficient number of studies with a common comparator (e.g., placebo). For all analyses, we will use random-effects models to estimate pooled or comparative effects.

To determine whether quantitative analyses are appropriate, we will assess the clinical and methodological heterogeneity of the studies under consideration following established guidance.[42] We will do this by qualitatively assessing the PICOTS of the included studies, looking for similarities and differences.

If we conduct quantitative syntheses (i.e., meta-analysis), we will assess statistical heterogeneity in effects between studies by calculating the chi-squared statistic and the $I^2$ statistic (the proportion of variation in study estimates attributable to heterogeneity). The importance of the observed value of $I^2$ depends on the magnitude and direction of effects and on the strength of evidence for heterogeneity (e.g., p-value from the chi-squared test, or a confidence interval for $I^2$). If we include any meta-analyses with considerable statistical heterogeneity in this report, we will provide an explanation for doing so, considering the magnitude and direction of effects. We will also examine potential sources of heterogeneity using sensitivity analysis or analysis of subgroups. We plan to stratify analyses and/or perform subgroup analyses when possible and appropriate to examine clinical heterogeneity.

For any quantitative analyses, we will conduct sensitivity analyses, including studies with a high risk of bias. The weight that we will give to studies with a high risk of bias will be based on the perceived potential bias in the results from each of these studies. Planned stratifications or categories for subgroup analyses include the subgroups listed in the

analytic framework and geographic location of studies. When quantitative analyses are not appropriate (e.g., because of heterogeneity, insufficient numbers of similar studies, insufficiency or variation in outcome reporting), we will synthesize the data qualitatively.

We will follow the Methods Guide[43] to integrate the systematic review evidence with our primary studies meeting inclusion criteria.

**Grading the Strength of Evidence for Individual Comparisons and Outcomes**: We will grade the strength of evidence based on the updated guidance in the Methods Guide. Developed to grade the overall strength of a body of evidence,[44] this approach incorporates five key domains: study limitations (includes study design and aggregate risk of bias), consistency, directness, precision of the evidence, and reporting bias. It also considers other optional domains that may be relevant for some scenarios, such as a dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect). These domains are particularly relevant for observational studies; if we expand our evidence base beyond trials, we will consider these domains in addition to the five key domains.

Grades reflect the strength of the body of evidence to answer KQs included in this review (see Table 2). Two reviewers will assess each domain for each key outcome, and differences will be resolved by consensus. Senior members of the review team (including at least one subject matter expert and one methodologist) will grade the strength of evidence for the outcomes deemed to be of greatest importance to decisionmakers (consulting with TEP members about this as needed) and those most commonly reported in the literature.

**Table 2. Definitions of the grades of overall strength of evidence[44]**

| Grade | Definition |
|---|---|
| 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:** We will assess applicability of the evidence following guidance from the Methods Guide.[45] We will use the PICOTS framework to explore factors that affect applicability. Some factors identified a priori that may limit the applicability of evidence include the following: healthcare delivery setting in system or organization having unique characteristics, differences in usual care characteristics of the system or organization or provider that comprise the comparison group, varying provider type, evidence-based practice type being implemented, intensity of the strategy, and timing of application of strategy.

## References


1. O'Connell ME, Boat T, Warner KE, eds. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. National Academies Press: Washington, DC; 2009.

2. Perou R, Bitsko RH, Blumberg SJ, et al. Mental health surveillance among children--United States, 2005-2011. MMWR Surveill Summ. 2013 May 17;62 Suppl 2:1-35. Epub: 2013/05/17. PMID: 23677130.

3. Silverman WK, Hinshaw SP. The second special issue on evidence-based psychosocial treatments for children and adolescents: a 10-year update. J Clin Child Adolesc Psychol. 2008 Jan-Mar;37(1):1-7.

4. Chorpita BF, Daleiden EL, Ebesutani C, et al. Evidence-based treatments for children and adolescents: an updated review of indicators of efficacy and effectiveness. Clin Psychol-Sci Pr. 2011 Jun;18(2):154-72. PMID: WOS:000291755900009.

5. Jensen PS, Goldman E, Offord D, et al. Overlooked and underserved: "action signs" for identifying children with unmet mental health needs. Pediatrics. 2011 Nov;128(5):970-9. Epub: 2011/10/26. PMID: 22025589.

6. Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: variation by ethnicity and insurance status. Am J Psychiatry. 2002 Sep;159(9):1548-55. Epub: 2002/08/31. PMID: 12202276.

7. Saloner B, Carson N, Cook BL. Episodes of mental health treatment among a nationally representative sample of children and adolescents. Med Care Res Rev. 2014 Jun;71(3):261-79. Epub: 2014/01/09. PMID: 24399817.

8. Satcher D. U.S. Public Health Service, Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Washington, DC: Department of Health and Human Services; 2008.

9. Williams J, Klinepeter K, Palmes G, et al. Diagnosis and treatment of behavioral health disorders in pediatric practice. Pediatrics. 2004 Sep;114(3):601-6. PMID: 15342827.

10. Weisz JR, Weiss B, Han SS, et al. Effects of psychotherapy with children and adolescents revisited: a meta-analysis of treatment outcome studies. Psychol Bull. 1995 May;117(3):450-68. Epub: 1995/05/01. PMID: 7777649.

11. Dishion TJ, McCord J, Poulin F. When interventions harm. Peer groups and problem behavior. Am Psychol. 1999 Sep;54(9):755-64. Epub: 1999/10/08. PMID: 10510665.

12. Hoagwood K, Burns BJ, Kiser L, et al. Evidence-based practice in child and adolescent mental health services. Psychiatr Serv. 2001 Sep;52(9):1179-89. Epub: 2001/09/05. PMID: 11533391.

13. Bickman L, Kelley SD, Breda C, et al. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. Psychiatr Serv. 2011 Dec;62(12):1423-9. Epub: 2011/12/24. PMID: 22193788.

14. Glisson C, Schoenwald SK, Hemmelgarn A, et al. Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. J Consult Clin Psychol. 2010 Aug;78(4):537-50. Epub: 2010/07/28. PMID: 20658810.

15. Balas EA, Boron SA. Managing clinical knowledege for healthcare improvements. Stuttgart, Germany: Schattauer Verlagsgesellschaft; 2000.

16. Shojania KG, McDonald KM, Wachter RM, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies, Vol. 1: Series Overview and Methodology. Technical Review 9 (Contract No. 290-02-0017 to the Stanford University-UCSF Evidence-based Practices Center) AHRQ Publication No. 04-0051-1. Rockville, MD: Agency for Healthcare Research and Quality; Aug 2004.

17. Dorsey S, Pullmann MD, Deblinger E, et al. Improving practice in community-based settings: a randomized trial of supervision - study protocol. Implement Sci. 2013;8:89. Epub: 2013/08/14. PMID: 23937766.

18. Schoenwald SK, Sheidow AJ, Letourneau EJ. Toward effective quality assurance in evidence-based practice: links between expert consultation, therapist fidelity, and child outcomes. J Clin Child Adolesc Psychol. 2004 Mar;33(1):94-104. Epub: 2004/03/19. PMID: 15028545.

19. McKibbon KA, Lokker C, Wilczynski NL, et al. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? Implement Sci. 2010 Feb 12;5(1):16. PMID: 21080976.

20. McCormack L, Sheridan S, Lewis M, et al. Communication and Dissemination Strategies to Facilitate the Use of Health-Related Evidence. Evidence Report/Technology Assessment No. 213. (Prepared by the RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13(14)-E003-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

21. Proctor EK, Landsverk J, Aarons G, et al. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health. 2009 Jan;36(1):24-34. Epub: 2008/12/24. PMID: 19104929.

22. Aarons GA, Glisson C, Green PD, et al. The organizational social context of mental health services and clinician attitudes toward evidence-based practice: a United States national study. Implement Sci. 2012;7:56. Epub: 2012/06/26. PMID: 22726759.

23. Aarons GA. Mental health provider attitudes toward adoption of evidence-based practice: the Evidence-Based Practice Attitude Scale (EBPAS). Ment Health Serv Res. 2004 Jun;6(2):61-74. Epub: 2004/07/01. PMID: 15224451.

24. Glisson C, Landsverk J, Schoenwald S, et al. Assessing the organizational social context (OSC) of mental health services: implications for research and practice. Adm Policy Ment Health. 2008 Mar;35(1-2):98-113. Epub: 2007/12/19. PMID: 18085434.

25. Mendel P, Meredith LS, Schoenbaum M, et al. Interventions in organizational and community context: a framework for building evidence on dissemination and implementation in health services research. Adm Policy Ment Health. 2008 Mar;35(1-2):21-37. Epub: 2007/11/09. PMID: 17990095.

26. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health. 2011 Jan;38(1):4-23. Epub: 2011/01/05. PMID: 21197565.

27. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:50. Epub: 2009/08/12. PMID: 19664226.

28. Meyers DC, Durlak JA, Wandersman A. The quality implementation framework: a synthesis of critical steps in the implementation process. Am J Community Psychol. 2012 Dec;50(3-4):462-80. Epub: 2012/05/31. PMID: 22644083.

29. Kaplan HC, Provost LP, Froehle CM, et al. The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. BMJ Qual Saf. 2012 Jan;21(1):13-20. PMID: 21835762.

30. Hoagwood K. Personal communication. Personal communications with; 2014.

31. Ringeisen H, Henderson K, Hoagwood K. Context matters: schools and the "research to practice gap" in children's mental health. School Psych Rev. 2003;32(2):153-68.

32. Rojas Smith L, Ashok M, Dy SM, et al. Contextual Frameworks for Research on the Implementation of Complex System Interventions. Methods Research Report. (Prepared by the RTI International-University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 14-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2014. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

33. Barwick MA, Schachter HM, Bennett LM, et al. Knowledge translation efforts in child and youth mental health: a systematic review. J Evid Based Soc Work. 2012;9(4):369-95. Epub: 2012/07/27. PMID: 22830938.

34. Novins DK, Green AE, Legha RK, et al. Dissemination and implementation of evidence-based practices for child and adolescent mental health: a systematic review. J Am Acad Child Adolesc Psychiatry. 2013 Oct;52(10):1009-25 e18. Epub: 2013/10/01. PMID: 24074468.

35. Powell BJ, McMillen JC, Proctor EK, et al. A compilation of strategies for implementing clinical innovations in health and mental health. Med Care Res Rev. 2012 Apr;69(2):123-57. PMID: 22203646.

36. McPheeters ML, Kripalani S, Peterson NB, et al. Quality Improvement Interventions To Address Health Disparities. Closing the Quality Gap: Revisiting the State of the Science. Evidence Report No. 208. (Prepared by the Vanderbilt University Evidence-based Practice Center under Contract No. 290-2007-10065.) AHRQ Publication No. 12-E009-EF. Rockville, MD: Agency for Healthcare Research and Quality; August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

37. White CM, Ip S, McPheeters M, et al. Using existing systematic reviews to replace de novo processes in conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews [posted September 2009]. Rockville, MD; 2009.

38. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews AHRQ Publication No. 12-EHC047-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2012. www.effectivehealthcare.ahrq.gov/.

39. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012 Feb;65(2):163-78. PMID: 21959223.

40. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10. PMID: 17302989.

41. 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):267-70. PMID: 24727843.

42. West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity. Methods Research Paper AHRQ Publication

No. 10-EHC070-EF. Rockville, MD: Agency for Healthcare Research and Quality; Sep 2010. http://effectivehealthcare.ahrq.gov/.

43. Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies. Draft Methods Guidance Rockville, MD: in process. http://effectivehealthcare.ahrq.gov/ehc/products/597/1964/Integrating-evidence-draft-140910.pdf.

44. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI International-University of North Carolina 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; November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

45. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions. Methods Guide for Comparative Effectiveness Reviews AHRQ Publication No. 11-EHC019-EF. Agency for Healthcare Research and Quality; January 2011. http://effectivehealthcare.ahrq.gov/.

46. Ragin CC. The comparative method: moving beyond qualitative and quantitative strategies. Berkeley: University of California Press; 1987.

47. Ragin CC. Using qualitative comparative analysis to study causal complexity. Health Serv Res. 1999;34(5 Pt 2):1225-39.

48. Ragin CC. Redesigning social inquiry: fuzzy sets and beyond. Chicago: University of Chicago Press; 2008.

49. Schneider CQ, Wagemann C. Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis. Cambridge: Cambridge University Press; 2012.

50. Kahwati L, Viswanathan M, Golin CE, et al. Identifying configurations of behavior change techniques in effective medication adherence interventions: a qualitative comparative analysis in review.

51. Kahwati L, Jacobs S, Kane H, et al. Using qualitative comparative analysis in a systematic review of a complex intervention in review.

52. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy. Oslo: Norwegian Knowledge Centre for the Health Services; 2002.

53. Berg-Schlosser D, De Meur G. Comparative research design: Case and variable selection. In: Rihoux Benoit RC, Ragin C, eds. In: Configurational comparative

methods: qualitative comparative analysis (QCA) and related techniques. Thousand Oaks, CA: SAGE Publications; 2009.

## V. Definition of Terms

Not applicable.

## VI. Summary of Protocol Amendments

| Date | Rationale | Original Protocol | Revised Protocol | Rationale |
|---|---|---|---|---|
| December 11, 2015 | Methods: Searching | | In addition to the described searches, we will search for publications related to studies included in the review authored by study investigators. This intended to identify more detail on the strategies reported to better understand the most important components of each strategy. We will also contact study investigators authors to obtain information about critical components for strategies of included studies. | All included studies investigated complex interventions. This information will contextualize the published data and inform additional analysis described below. |
| December 11, 2015 | Methods: Assessment of Methodological Risk of Bias of Individual Studies | | We will not assess study quality of publications related to included studies. | These related publications are intended to provide context and additional detail about the strategies used in the included studies. |
| December 11, 2015 | Methods: Data Synthesis | Quantitative methods will be used if possible, otherwise qualitative methods will be used to synthesize the data. | Two sets of additional analyses will be used to synthesize and contextualize the data. First, we will reach out to authors of included studies to describe the critical components of the strategies they tested so that we can evaluate them for common elements. Second, we will search for other publications by authors of included studies that provide more details on the strategies reported to | These additional methods are planned because the strategies were all complex in nature. These methods are intended to help improve the synthesis of findings by identifying critical components of the strategy to best improve outcomes. |

| Date | Rationale | Original Protocol | Revised Protocol | Rationale |
|---|---|---|---|---|
| | | | better understand the most important components of each strategy. This will enable us to provide greater clarity in the report regarding the driving forces of each strategy. Additional detail will be added to the description of the strategy. This will also allow us to identify critical components of a strategy in relation to its outcomes. | |
| December 11, 2015 | Methods: Data Synthesis | Quantitative methods will be used if possible, otherwise qualitative methods will be used to synthesize the data. | In addition to the planned analysis, we will use quantitative comparative analysis (QCA) using the dataset abstracted for the primary synthesis to examine set relationships between combinations of strategy components and improvements in outcomes. For detailed methods please see Appendix B. | This additional method is planned because the strategies were complex in nature and heterogeneous. While hypothesis-generating, this analysis may provide information about combinations of strategy components or factors more likely to be effective. |

## VII. Review of Key Questions

For all EPC reviews, KQs were reviewed and refined as needed by the EPC with input from Key Informants and the TEP to ensure that the questions are specific and explicit about what information is being reviewed. In addition, the KQs were posted for public comment and put into final form by the EPC after review of the comments.

## VIII. 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 health care 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 Task Order Officer and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

## IX. Technical Experts

Technical Experts constitute a multidisciplinary group of clinical, content, and methodologic 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 conflicted opinions are common and perceived as healthy 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 or 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.

## X. Peer Reviewers

Peer Reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. In preparing the final draft of the report, the EPC considers all peer review comments on the preliminary draft. Peer Reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for comparative effectiveness reviews and technical briefs, be published 3 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.

## XI. 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 total more than $1,000 will usually disqualify EPC core team investigators.

## XII. Role of the Funder

This project was funded under Contract No. HHSA290-2012-00008-I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

## Appendix A: Preliminary Literature Search Terms for MEDLINE®

| Search | Query |
|---|---|
| #1 | Search "Health Plan Implementation"[Mesh] |
| #2 | Search ("Quality Improvement"[Mesh] OR "quality improvement"[All Fields] OR "quality initiative"[All Fields]) |
| #3 | Search (("Information Dissemination"[Mesh] OR "Diffusion of Innovation"[Mesh] OR "Health Information Management"[Mesh])) |
| #4 | Search (#1 or #2 or #3) |
| #5 | Search (("Patient Acceptance of Health Care"[Mesh] OR adaptation[tiab] OR disseminat*[tiab] OR "Feasibility Studies"[Mesh] OR feasibility[tiab] OR fidelity[tiab] OR implement*[tiab] OR penetration[tiab] OR supervision[tiab] OR sustain*[tiab] OR "Information Systems"[MeSH] OR uptake[tiab])) |
| #6 | Search (("Guideline Adherence"[Mesh] OR "Evidence-Based Practice"[Mesh] OR "evidence based practice"[All Fields] OR "evidence-based practice"[All Fields] OR effect* OR evidence)) |
| #7 | Search (#5 and #6) |
| #8 | Search (#4 or #7) |
| #9 | Search ("mental health"[All Fields] OR "mental illness"[All Fields] OR "mental disorders"[All Fields] OR "mental disorder"[All Fields] OR psychopathology OR "Adjustment Disorders"[Mesh] OR "adjustment disorder"[All Fields] OR "anxiety disorder"[All Fields] OR agoraphobia OR "panic disorder"[All Fields] OR "Phobic Disorders"[Mesh] OR phobia OR "Stress Disorders, Post-Traumatic"[Mesh] OR "posttraumatic stress disorder"[All Fields] OR "post-traumatic stress disorder"[All Fields] OR "generalized anxiety disorder"[All Fields] OR "Obsessive-Compulsive Disorder"[Mesh] OR "obsessive compulsive disorder"[All Fields] OR "reactive attachment disorder"[All Fields] OR "Anxiety, Separation"[Mesh] OR "separation anxiety disorder"[All Fields] OR "Eating Disorders"[Mesh] OR "eating disorder"[All Fields] OR "anorexia nervosa"[All Fields] OR "bulimia nervosa"[All Fields] OR "Attention Deficit Disorder with Hyperactivity"[Mesh] OR "attention deficit hyperactivity disorder"[All Fields] OR "Attention Deficit and Disruptive Behavior Disorders"[Mesh] OR "conduct disorder"[All Fields] OR "oppositional defiant disorder"[All Fields] OR depression OR "depressive disorder"[All Fields] OR "Bipolar Disorder"[Mesh] OR "bipolar disorder"[All Fields] OR mania OR "dysthymic disorder"[All Fields] OR "Schizophrenia"[Mesh] OR schizophrenia OR "Psychotic Disorders"[Mesh] OR "psychotic disorder"[All Fields] OR encopresis OR "Personality Disorders"[Mesh] OR "personality disorder"[All Fields] OR "behavioral disorder"[All Fields] OR "behavioral disturbance"[All Fields] OR "serious emotional distress"[All Fields] OR "emotional disorder"[All Fields] OR "Substance-Related Disorders"[All Fields] OR "substance use disorder"[All Fields] OR "drug use disorder"[All Fields] OR "Alcohol-Related Disorders"[Mesh] OR "alcohol use disorder"[All Fields] OR "alcohol dependence"[All Fields] OR alcoholism OR "drug dependence"[All Fields] OR "cannabis dependence"[All Fields] OR "marijuana dependence"[All Fields] OR "Tobacco Use Disorder"[Mesh] OR "nicotine dependence"[All Fields] OR "substance dependence"[All Fields] OR "substance abuse"[All Fields] OR "alcohol abuse"[All Fields] OR "drug abuse"[All Fields] OR "cannabis abuse"[All Fields] OR "marijuana abuse"[All Fields]) |
| #10 | Search (#8 and #9) |
| #11 | Search (("diffusion tensor" OR "diffusion tensors")) |
| #12 | Search (#10 not #11) |
| #13 | Search (#10 not #11) Filters: Editorial |
| #14 | Search (#10 not #11) Filters: Editorial; Letter |
| #15 | Search (#12 NOT #14) |
| #16 | Search (((randomized[title/abstract] AND controlled[title/abstract] AND trial[title/abstract]) OR (controlled[title/abstract] AND trial[title/abstract]) OR "controlled clinical trial"[publication type] OR "Randomized Controlled Trial"[Publication Type] OR "Single-Blind Method"[MeSH] OR "Double-Blind Method"[MeSH] OR "Random Allocation"[MeSH])) |
| #17 | Search (#15 and #16) |
| #18 | Search (("Cohort Studies"[MeSH] OR (prospective AND cohort))) |
| #19 | Search (#15 and #18) |
| #20 | Search (#12 NOT #14) Filters: Review |

| Search | Query |
|---|---|
| #21 | Search (#17 or #19 or #20) |
| #22 | Search (#17 or #19 or #20) Filters: Child: birth-18 years |

## Appendix B

Qualitative comparative analysis (QCA) is a theory-driven approach that is particularly suited to understanding complex causal patterns within and across cases. QCA uses formal logic, a branch of mathematics, to examine combinations of conditions (in our study, components of a complex intervention) and their relationship to an outcome. Individual conditions and combinations of conditions can be necessary, sufficient, or both, to the outcome. Necessary conditions (and combinations of conditions) are antecedent to the outcome: the outcome cannot occur in the absence of the necessary condition. Sufficient conditions (and combinations of condition) guarantee the outcome. For example, access to evidence-based interventions is necessary for uptake of evidence-based intervention, but it may not be sufficient for uptake; other variables may be necessary as well (e.g., administrative support for clinical staff).

QCA accommodates both qualitative data and quantitative data within the same analysis, and offers a systematic way for evaluating causal complexity because it is based on formal logic and set theory, not statistical theory. Traditional variable-oriented methods typically deconstruct the unit of analysis into its component variables and then assess statistical correlations among one or more variables, but this may not be the best approach for complex interventions. QCA can identify multiple "recipes" or sufficient combinations for achieving an outcome; in other words, one size does not fit all. For example, in Kahwati et al.'s application of QCA to a systematic review on medication adherence the authors found several *combinations of behavioral techniques* in included studies that led to improved adherence.

Several authors describe QCA methods in greater detail.[46-51] We base our methods on the approach used by Kahwati et al. in a recent QCA of medication adherence studies.[50,51] Because QCA requires iterative rounds of analysis, we anticipate that some of the methods described below may evolve. We will record all planned and conducted analysis to ensure transparency. We plan to include as many of studies in this review as possible, but we recognize this may not be feasible, or always appropriate.

1. **Specify the configural questions.** For this review, we ask "what combination of intervention components are present in studies demonstrating improved implementation, dissemination, quality improvement?"
2. **Identify cases for use in analysis**. All included studies had at least two arms but did not always provide information on arm-specific improvements. Therefore, we cannot include each arm in each study as a case, rather, each case constitutes a comparison between two study arms. A study with three arms provides 2 cases for analysis.
3. **Specify and calibrate condition sets**. In this step, we plan to examine several causal conditions. Specifically, we focus on intervention components as defined by the taxonomy used by the international Cochrane Review Group's Effective

Practice and Organisation of Care (EPOC) Group to classify complex strategies designed to improve health care professionals' practice and the organization of health care services.[52] We classify whether study arms differ with respect to at least one professional component (e.g., distribution of educational materials, educational meetings, audit and feedback, etc.), at least one financial component (both provider financial components such as provider incentives and provider grants/allowances and patient financial components such as patient incentives, and at least one organizational component (both provider-oriented organizational components such as clinical multidisciplinary teams and provider satisfaction as well as structural organizational components such as changes in scope and nature of benefits and services, staff organization, presence and organization of quality monitoring mechanisms, etc.)

In QCA, condition (and outcome sets) are calibrated by establishing thresholds and decision rules for membership in a condition. In a crisp set, a value of 1 indicates that a case is fully in the condition set; a value of 0 indicates that a case is fully out of a condition set. In keeping with the definition of cases as comparisons between two arms, we define conditions as differences between two arms. We use the EPOC taxonomy to classify main intervention components, dually and independently, and will resolve conflicts after consensus. A value of 0 indicates there was no difference between the arms in that component. A value of 1 means that the arms varied in that component. For example, cases assigned a condition value of 1 for financial components have at least one financial component (such as pay-for-performance) that the control arm did not have.

4. **Specify and calibrate outcome set**. We intend to assess "having evidence of improvement" as our outcome set. As with calibration of the condition set, we will dually and independently calibrate the outcome set, and resolve conflicts after consensus. We assign a value of 1, indicating evidence of improvement) to cases that demonstrated a statistically significant difference between arms for at least one measure of intermediate practitioner, systems, or patient outcomes. We assign a value of 0, indicating no evidence of improvement, to cases that demonstrated a statistically significant difference between arms for at least one measure of intermediate practitioner, systems, or patient outcomes.

5. **Construct and analyze the truth table.** The truth table, the key analytic device in QCA, helps determine which combinations of conditions occur consistently with improvement. We will use fsQCA Version 2.5 to identify solutions (i.e., combinations conditions that are necessary or sufficient for the outcomes). This analysis also will include examination of parameters of fit: consistency and coverage. *Consistency* assesses whether the causal pathway produces the outcome

regularly ("the degree to which the empirical data are in line with a postulated subset relation," Ragin, 1999, p. 324); *coverage* determines the empirical relevance of a solution and quantifies the variation in causal pathways to an outcome (analogous to variance explained, or how well the solutions explain outcomes across all included cases). The results of a QCA analysis are statements of necessity and sufficiency, expressed as text, solution formulas, or in tabular or graphic formats.

6. **Make sense of the results**. We will return to the included studies to evaluate the identified solutions and understand the contextual elements that might explain these solutions.

Our proposed analyses will acknowledge several limitations of QCA. First, models can only be used to investigate a few conditions of interest because QCA examines each possible combination of conditions, which exponentially increases with each addition (i.e., 5 conditions=32 possible combinations, 6 conditions=64 possible conditions). Thus, combinations that have no data cannot be analyzed. Another limitation is that the strategies themselves will need to be at least somewhat comparable,[53] allowing for investigation of only high-level components consistent across strategies.