Comparative Effectiveness Review
Number 181

Effective Health Care Program

Strategies To Improve Mental Health Care for Children and Adolescents

Executive Summary

Background

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. They may exacerbate coexisting conditions and may reduce quality of life. They also increase the risk of involvement with the criminal justice system and other risk-taking behaviors and suicide.[2]

Several key publications in the mid- to late 1990s suggested that usual care in children's mental health had, at best, no[3] and sometimes harmful effects.[4] Since then, mental health interventions that improve children and adolescents with mood disorders, anxiety disorders, disruptive behavior disorders, psychotic disorders, eating disorders, and substance use disorders have been tested to varying degrees of benefit.[5,6]

Despite advances in the evidence base,[5,7] some outcomes for children with mental health problems remain suboptimal because of issues with access to care and the failure of systems and providers to adopt established quality improvement (QI) strategies and interventions with proven effectiveness (e.g., evidence-based practices [EBPs]). 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

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm**.

AHRQ Agency for Healthcare Research and Quality
Effective Health Care Program

treatment episodes are considered minimally adequate (at least four visits with psychotropic medication or at least eight visits without psychotropic medication).[8-10] The current health care system continues to provide fragmented care to children and adolescents in numerous uncoordinated systems, rendering inefficient the delivery of needed services.[11] Moreover, clinicians—particularly primary care practitioners—may lack the time, knowledge, or training to identify and treat or refer patients with mental health problems.[12]

Given the gap between observed and achievable processes and outcomes, one way to improve the mental health care of children and adolescents is to adopt QI strategies and develop strategies to implement or disseminate interventions with known effectiveness. Such strategies target changes in the organization and delivery of mental health services.[13,14] They seek to improve the quality of care and patient outcomes by closing the gap between research evidence and practice.[15-17]

The ultimate goal of these strategies is to improve patient health and service utilization outcomes for children and adolescents with mental health problems. Intermediate outcomes in this context include changes to health care systems, organizations, and practitioners that provide mental health care. Targeting multiple, interrelated, nested levels such as the macro environment (e.g., state), organization or system (e.g., specialty mental health clinic), program (e.g., selected intervention), practitioners (e.g., clinicians), and patients (e.g., children or adolescents and their families) typically increases the effectiveness and sustainability of a particular strategy.[18,19] For instance, changes in intermediate outcomes such as practitioners' attitudes[20] or organizational climate[21] may influence the successful adoption of and fidelity to EBPs. These practices in turn influence patient health outcomes, such as behavior or quality of life.

Scope and Key Questions

Key Questions (KQs)

KQ 1: What is the effectiveness of QI, implementation, and dissemination strategies employed in outpatient settings by health care practitioners, organizations, or systems that care for children and adolescents with mental health problems to improve:

a. intermediate patient, provider, or system outcomes

b. patient health and service utilization outcomes?

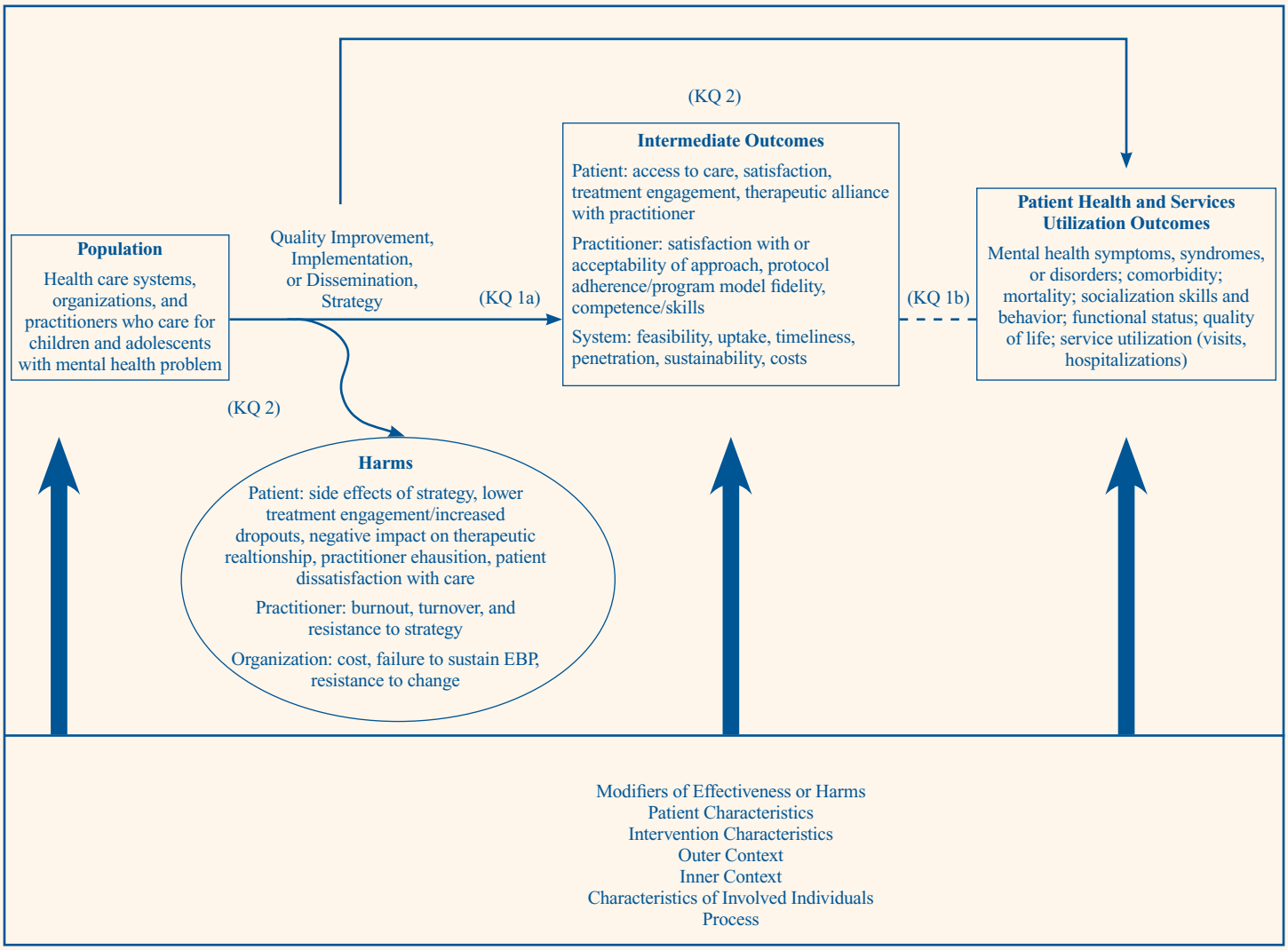
KQ 2: What are the harms of these mental health strategies?

KQ 3: Do characteristics of the child or adolescent or contextual factors (e.g., characteristics of patients, practitioners, organizations, or systems; intervention characteristics; setting; or process) modify the effectiveness or harms of strategies to improve mental health care and, if so, how?

Analytic Framework

Figure A depicts the patient populations, interventions, comparators, outcomes, and timing of outcomes assessment (PICOTs) and KQs in relation to these PICOTs.

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



EBP = evidence-based practices; KQ = Key Question.

Populations, Interventions, Comparators, Outcomes, Timing, and Setting

We specified our inclusion and exclusion criteria based on the PICOTS early in the systematic review process after conducting a literature scan and receiving input from key informants. We included QI, implementation, and dissemination strategies that targeted systems, organizations, or practitioners of mental health care to children and adolescents 18 years of age or younger, who were already experiencing mental health symptoms. As a result, universal interventions aimed at prevention are not included. We did not include strategies such as the implementation of educational interventions for reading disorders. We also limited our review of implementation strategies to those focusing on EBP interventions. For defining EBPs, we relied on the minimum requirements set forth by the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-based Programs and Practices (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.

We use the term "strategy" to reference the total sum of components used to target health care systems and/or practitioners to improve the quality of care for children and adolescents with mental health problems. We use the term "intervention" to denote a specific EBP used as part of a strategy.

Because strategies tended to be complex in nature and the number and types of components that varied between the treatment arm and comparison group arm differed by study, we also recorded components of each strategy. We relied on the Cochrane Review Group's Effective Practice and Organisation of Care (EPOC) Group taxonomy, which categorizes strategies by whether they include one or more professional, financial, organizational, and regulatory components.[22] Because many of the comparison groups also contained several components, we marked the components contained in each study arm of each study. This allowed us to fully describe the numerous components that were being combined and tested in each strategy, as well as enabled us to determine whether the study arms differed by a single or multiple components.

We required each included study to report at least one intermediate outcome in a minimum of one of three major categories: (1) practitioner intermediate outcomes (satisfaction, adherence, fidelity, competence), (2) system intermediate outcomes (feasibility, uptake, timeliness, penetration, sustainability, costs), and (3) patient intermediate outcomes (access to care, satisfaction, engagement, therapeutic alliance). This approach helped ensure that each included study demonstrated impact based on its stated goals of improving quality or implementing or disseminating evidence-based interventions. We also required each study to report at least one patient health or service utilization outcome (change in mental health status, comorbid conditions, mortality, socialization skills and behavior, functional status, quality of life, service utilization) if the strategy was not implementing or disseminating an EBP (i.e., an intervention with proven effectiveness).

For all KQs, we excluded study designs without comparison groups to ensure that our pool of included studies provided strong evidence on the causal link between the strategy and outcomes. We also required that the comparator enabled examination of the strategy effectiveness. That is, we excluded studies in which the strategy (system, organizational, practitioner targets) and the intervention being tested both differed between groups, because the effectiveness of the QI, implementation, or dissemination strategy could not be isolated from the baseline intervention effects.

Our exclusion of non-English-language studies is based on limitations of time and resources. However, we examined English language abstracts of non-English-language studies to assess the potential size of the literature that would be missed through this approach.

Methods

The methods for this systematic review follow the *Methods Guide for Effectiveness and Comparative Effectiveness Review*s from Agency for Healthcare Research and Quality (AHRQ) (available at http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm). The review uses the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist to facilitate the preparation and reporting of the systematic review.[23]

Topic Refinement and Protocol Review

We developed this topic and KQs through a public process. AHRQ nominated the topic and we developed and refined it. Initially, a panel of Key Informants gave input on the KQs to be examined; AHRQ then posted these questions on the Effective Health Care Website for public comment from September 15, 2014, through October 6, 2014. We revised the KQs in response to comments.

We then drafted a protocol for the systematic review and recruited a panel of technical experts to provide high-level content and methodological expertise throughout the development of the review. The final protocol was posted on the Effective Health Care website at http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2030 on December 30, 2014, and registered on PROSPERO (Registration number: CRD42015024759). Following release of our draft report and peer review, we amended our protocol to include additional review and analysis strategies suitable for complex interventions (described under "Data Synthesis").

Literature Search Strategy

We systematically searched, reviewed, and analyzed the scientific evidence for each of our three KQs. We began 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 (from inception through January 14, 2016). We also searched the Cochrane Library, PsycINFO®, and CINAHL® (Cumulative Index to Nursing and Allied Health Literature) using analogous search terms.

In addition, we searched the gray literature (information that is unpublished and not controlled commercially) for studies relevant to this review and included studies that met all the inclusion criteria and contain enough methodological information to assess risk of bias. Sources of gray literature include ClinicalTrials.gov, the World Health Organization's International Clinical Trials Registry Platform, the National Institutes of Health Research Portfolio Online Reporting Tools, the Database of Promoting Health Effectiveness Reviews, and CMS.gov. To avoid retrieval bias, we manually searched 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.

Trained reviewers abstracted important information from included studies into evidence tables, housed on AHRQ's Systematic Review Data Repository. A second senior member of the team reviewed all data abstractions for completeness and accuracy. Reviewers resolved conflicts by discussion and consensus or by consulting a third member of the review team.

Risk of Bias Assessment

To assess the risk of bias (internal validity) of studies, two independent reviewers used predefined, design-specific criteria based on guidance in the *Methods Guide*.[24] We resolved conflicts by consensus or by consulting a third member of the team. For randomized controlled trials (RCTs), we relied on the risk of bias tool developed by the Cochrane Collaboration.[25] We assessed the risk of bias of observational studies using questions from an item bank developed by RTI International[26] and A Cochrane Risk Of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI).[27] Minimum eligibility criteria for systematic reviews included an explicit description of search strategy used and determination that the search strategy was adequate, application of predefined eligibility criteria and risk of bias assessment for all included studies, and synthesis of the results presented.

In general terms, a study with no identifiable flaws has a low risk of bias. 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 (stemming from, for example, serious errors in design or conduct) that may invalidate its results. We considered the risk of bias for each relevant outcome of a study. When studies did not report sufficient detail to assess the validity of the design or study conduct, we judged the risk of bias to be unclear.

Data Synthesis

To determine whether quantitative analyses were appropriate, we assessed the clinical and methodological heterogeneity of the studies under consideration following established guidance.[28] For all outcomes, we present relative risks or mean differences, with confidence intervals (CIs), whenever calculable. For outcomes with multiple measures, we present forest plots.

We employed several other methods to provide additional information about the nature of the strategies tested and what components of the strategies had the most impact on outcomes. First, we performed additional search approaches of related publications (known as "cluster searching") to identify sibling (multiple publications on the same study) or kinship studies (publications from a common antecedent study or common theoretical foundation).[29] We hoped to uncover contextual information to explain failure or success of strategies. We also contacted study authors to obtain information about critical components for strategies of included studies as part of a parallel project to better understand the uses and limitations of trial registries for data on outcomes. This effort provided additional information on the important components of the strategies tested in included studies. Finally, we used qualitative comparative analysis (QCA) to examine set relationships between combinations of strategy components to identify those that were most associated with improvements in outcomes.

Strength of the Body of Evidence

We graded the strength of a body of evidence based on the updated guidance in the *Methods Guide*.[30,31] The AHRQ EPC approach incorporates five key domains: study limitations, 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.

Two reviewers assessed each domain for each key outcome and resolved any differences by consensus discussion. Senior members of the review team graded the strength of evidence.

Grades reflect the confidence that the reviewers have that various estimates of effect are close to true effects with respect to the KQs in a systematic review. Table A defines the four grades.

Table A. Definitions of the grades of overall strength of evidence[30]

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.

Risk of bias assessments for individual studies feed into the rating for the first of the strength of evidence domains, study limitations. Specifically, we rated bodies of evidence comprising trials with a high risk of bias as having high study limitations. Medium or unclear risk of bias studies resulted in medium study limitations. Low risk of bias studies resulted in low study limitations. In keeping with GRADE and strength of evidence guidance, we rated observational studies as having high study limitations.[31,32]

As described above, study design and study limitations together set the baseline strength of evidence grade. Other domains then could either reduce or increase the grade. A body of evidence with high study limitations, with no other reasons to increase confidence (dose-response, large magnitude of effect, plausible confounding) or decrease it (inconsistency, imprecision, indirectness, reporting bias) would generally have a low strength of evidence grade. A body of evidence with low study limitations, with no reasons to decrease confidence (inconsistency, imprecision, indirectness, reporting bias), would generally have a high strength of evidence grade. In other words, although study design and study limitations provide a baseline judgment of strength of evidence, each of four additional sources of uncertainty (inconsistency, imprecision, indirectness, reporting bias) serve to further reduce the strength of evidence grade.

For each source of uncertainty, we consistently used the following rubric to evaluate its effect on the overall strength of evidence across outcomes. Specifically, for indirectness, we rated intermediate outcomes as direct, rather than indirect, evidence. For this systematic review, these outcomes can be interpreted as direct measures of process change. Regarding consistency, we rated it as unknown for bodies of evidence with single studies; the rating of unknown consistency did not lower the overall grade. We relied on established guidance to judge precision.[33] Regarding imprecision, we specified the reasons for our judgment (small sample size or event rate, particularly when considering the optimum information size for the specific outcome, CIs crossing the line of no difference, or very wide CIs).[32] We downgraded the overall strength of evidence by two levels when we found multiple reasons for imprecision. We upgraded the evidence by one level for factors such as large magnitude of effect.

Applicability

We assessed applicability of the evidence following guidance from the *Methods Guide*.[34] We used the PICOTS framework to explore factors that affect applicability.

Results

We provide a summary of results by KQ below. Detailed descriptions of included studies, key points, detailed synthesis, summary tables, and expanded strength of evidence tables that include the magnitude of effect can be found in the full report. Our summary of results below presents the strength of evidence grades.

Results of Literature Searches

Figure B presents our literature search results through January 14, 2016. We found 17 eligible articles representing 17 studies[13,14,35-49] (one article reports on

two different studies[44] and another two articles[40,49] report outcomes for the same trial). We did not find any relevant non-English studies with English abstracts upon review.

This evidence base for KQ 1 consisted of 17 studies.[13,14,35-49] One of these studies addressed KQ 2 (harms) and four addressed KQ 3 (moderators of effectiveness). The evidence base included RCTs,[13,14,35-37,39,40,42,44-49] controlled clinical trials (CCTs),[41,43] interrupted time series,[38] and cohort designs.[44] Full evidence tables are available at http://srdr.ahrq.gov/projects/530.

We classified strategies with one or more financial or organizational components as "financial or organizational change" strategies and strategies with only professional components as "professional training" strategies. These categories guided our qualitative synthesis. We present summary tables of descriptions of strategy components and differences by study arms for each included study in the text of our main report. Table B presents study characteristics for professional training and financial or organizational change strategies.

Figure B. Results of literature searches

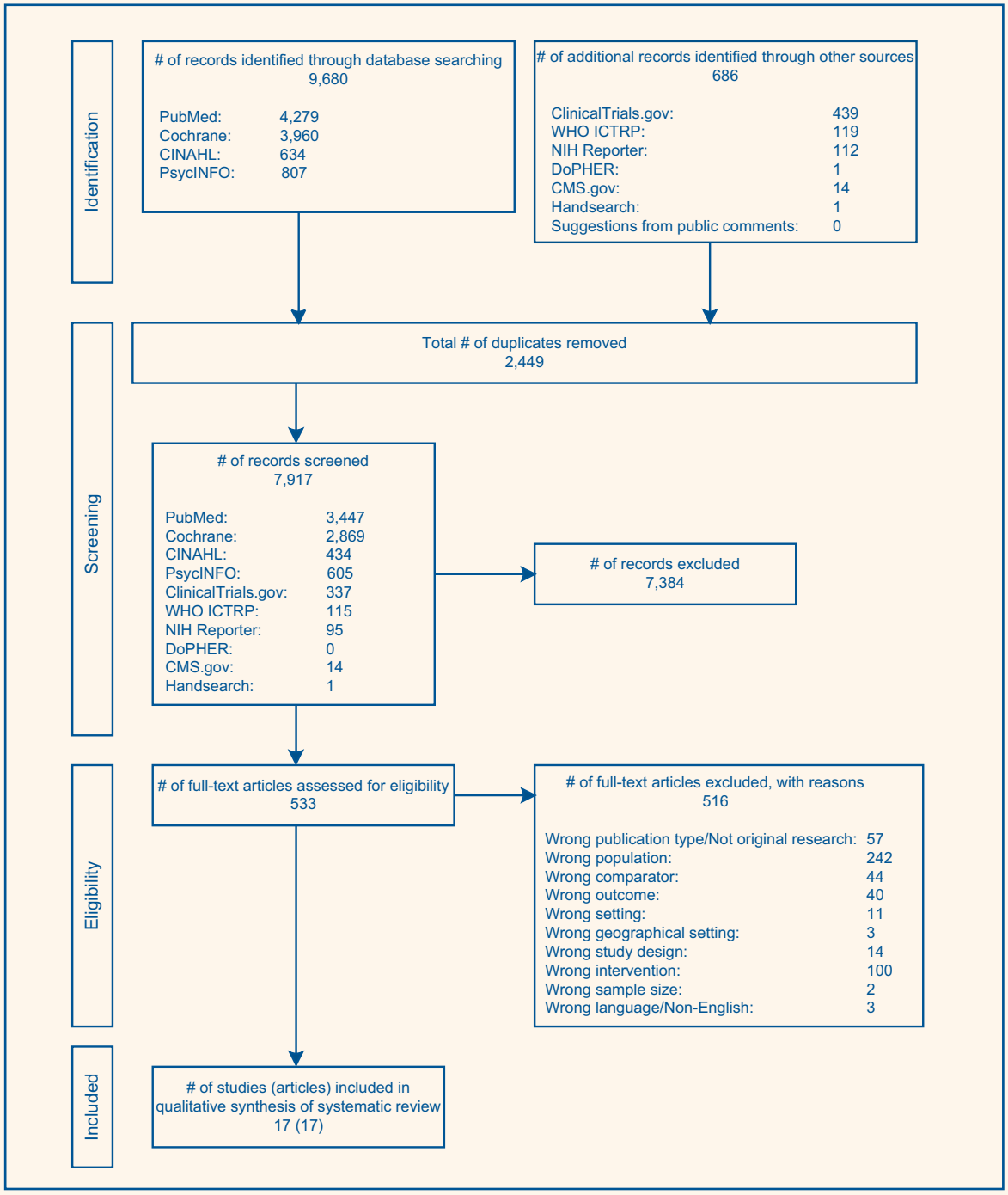


Table B. Strategies to improve mental health of children and adolescents: Study characteristics

Study Descriptor	Characteristics	Primary Strategy: Professional Training[a]	Primary Strategy: Financial or Organizational Change[b]	Total
Design	RCT	2	0	2
	2-stage RCT	0	1	1
	Cluster RCT	3	7	10
	CCT	0	2	2
	Non-RCT	2	0	2
Setting	Primary care	1	2	3
	Community mental health	4	8	12
	School	1	0	1
Strategy Categorization[c]	Quality improvement	2	3	5
	Implementation	1	4	5
	Dissemination	0	0	0
	Hybrid QI and I	1	2	3
	Hybrid QI and D	2	1	3
	Hybrid I and D	1	0	1
Risk of Bias	Low	1	0	1
	Medium	0	2	3
	High	3	3	6
	Unclear	3	3	7
Key Question	KQ 1	7	10	17
	KQ 2	1	0	1
	KQ 3	1	3	4
Total N of studies		7	10	17

[a] Included all professional components from the EPOC taxonomy

[b] Included at least 1 financial or organizational component from the EPOC taxonomy

[c] Categories dually assigned by members of the study team according to the definitions of QI, I, and D included in the PICOTS

CCT = controlled clinical trial; D = dissemination; I = implementation; KQ = Key Question; N = number; QI = quality improvement; RCT = randomized controlled trial.

Below, we summarize the main findings. We then discuss the findings in relationship to what is already known, applicability of the findings, implications for decisionmaking, limitations, research gaps, and conclusions.

Key Findings and Strength of Evidence

Key Question 1. Effectiveness of Strategies To Improve Mental Health Care for Children and Adolescents

Table C describes interventions and summarizes the evidence for included studies. Most strategies were complex and included multiple (two to seven) different components (as defined by the EPOC taxonomy). We graded the strength of evidence of 28 outcomes for professional training strategies and of 19 for financial or organizational change strategies.

Table C. Strategies to improve mental health of children and adolescents: Summary table

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Adding an active learning component to a professional training workshop to implement an EBP	Cluster RCT, 115 therapists	Anxiety Ages 8–17 years	Augmented active learning vs. routine professional training workshop	Educational meetings or materials	No differences between arms for practitioner satisfaction with approach, protocol adherence, or practitioner skill	Low for no benefit for practitioner satisfaction, adherence, and skill	Low risk of bias, small sample size, imprecise results
Beidas et al., 2012[39]			Computerized routine training vs. routine professional training workshop	Educational meetings or materials	No differences between arms for practitioner protocol adherence or program model fidelity, or skill; computerized training group practitioners less satisfied than routine training group practitioners	Low for no benefit for practitioner satisfaction, adherence, and skill	Low risk of bias, small sample size, imprecise results
Adding weekly feedback to practitioners regarding patient symptoms to practitioners	Cluster RCT, N of clinicians unclear, 340 youth, 144 clinicians, 383 caregivers	General mental health problem (children who receive home-based mental health treatment) Mean age = 15 years	Weekly and cumulative 90-day feedback vs. cumulative 90-day feedback only on patient symptoms and functioning to practitioners	Audit and feedback	Two-thirds of practitioners did not view Web module	Insufficient for practitioner adherence	High study limitations, unknown precision for adherence
Bickman et al., 2011[13]					Membership in the weekly feedback group increased the rate of decline in functional severity scale by 0.01 (range: 1 to 5, higher scores indicate greater severity)	Low for benefit for functional severity	High study limitations, precise results for symptoms

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Adding diagnosis and treatment guidelines to a computer decision support system Carroll et al., 2013[35]	Cluster RCT, 84 patients	General mental health problem (children who receive home-based mental health treatment) Mean age = 15 years	Computer decision support plus electronic health record (EHR) that included diagnosis and treatment guidelines vs. computer decision support plus EHR only	Educational meetings or materials Patient-reported data Reminders Quality monitoring	Practitioner adherence improved through uptake of guidelines for diagnostic assessment (aOR, 8.0; 95% CI, 1.6 to 40.6); more reporting of 3 of 4 symptom domains at diagnosis	Low for benefit for practitioner adherence and program model fidelity	Medium study limitations, imprecise results with small number of events, large magnitude of effect
					No statistically significant differences on practitioner adherence through reassessment of symptoms at 3 months, adjustment of medications, and mental health referral	Insufficient for practitioner adherence (reassessment of symptoms) at 3 months, adjustment of medications, and referral	Medium study limitations, imprecise results (CIs cross the line of no difference)
					Visit to a mental health specialist calculated OR: 2.195; 95% CI, 0.909 to 5.303; p=0.081; reported p-value in study=0.054	Insufficient for service utilization	Medium study limitations, imprecise results (CIs cross the line of no difference)
Providing practitioner access to practice guidelines via an Internet portal Epstein et al., 2011[45]	Cluster RCT, 746 patients	Attention deficit hyperactivity disorder (ADHD) Ages 6 to 12 years	Internet portal providing practitioner access to practice guidelines vs. wait-list control	Educational meetings or materials Patient-reported data Audit and feedback Reminders Quality monitoring	Strategy appeared to improve 4 of 5 examined outcomes that measured practitioner protocol adherence and program model fidelity outcomes (mean change in proportion of patients who received targeted, evidence-based ADHD care outcomes between groups ranged from 16.6 to -50), but estimates were very imprecise, with large CIs	Low for benefit for practitioner protocol adherence and program model fidelity	Medium study limitations, imprecise (wide CIs)

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Collaborative consultation to promote the use of titration trials and periodic monitoring during medication management Epstein et al., 2007[36]	Cluster RCT, 38 practitioners, 144 patients	ADHD Mean age = 7 years	Collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management vs. control	Audit and feedback Multidisciplinary team	Practitioner adherence/ fidelity as measured by use of titration trials β=-0.283; SE, 0.09; p<0.01 and by use of medication monitoring trials: p=NS, details NR	Insufficient for practitioner adherence and fidelity	High study limitations, imprecise results (small sample size)
					Lower odds with overlapping confidence intervals of practitioner citing obstacles to implementation of EBP in 6 of 8 measures (2 reached statistical significance)	Insufficient for practitioner competence/skills	High study limitations, imprecise results (small sample size)
					F score for decrease in combined parent and teacher ratings of ADHD symptoms for group x time interaction: $F_{2, 144}$ = 0.44, p=0.65	Insufficient for patient change in mental health symptoms	High study limitations, imprecise results (small sample size)

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Paying practitioners for performance in implementing an EBP Garner et al., 2012[42]	Cluster RCT, 49 therapists, 936 patients	Substance use disorders Mean age = 16 years	Paying practitioners for performance in successfully delivering an EBP intervention vs. implementation as usual	Provider incentives	Therapists in the P4P group were over twice as likely to demonstrate implementation competence compared with IAU therapists (Event Rate Ratio, 2.24; 95% CI, 1.12 to 4.48)	Moderate for benefit for practitioner competence	Medium study limitations, precise results
					Patients in the P4P condition were more than 5 times as likely to meet target implementation standards (i.e., to receive specific numbers of treatment procedures and sessions) than IAU patients (OR, 5.19; 95% CI, 1.53 to 17.62)	Low for benefit for practitioner adherence and program fidelity	Medium study limitations, imprecise results (wide CIs)
					No statistically significant differences between groups OR, 0.68; 95% CI, 0.35 to 1.33	Low for no benefit for patient change in mental health symptoms	Medium study limitations, precise results
Program to improve organizational climate and culture Glisson et al., 2010[14 a]	Two-stage RCT, 596 youth, 257 therapists	Externalizing behaviors (youth referred to juvenile court with behavioral or psychiatric symptoms that require intervention) Ages 9–17 years	Program to improve organizational climate and culture vs. usual care	Educational meetings or materials Educational outreach visits Provider satisfaction initiative Audit and feedback	Details NR but does not demonstrate improvements in any measure of adherence by strategy group for any ARC vs. no ARC comparison	Low for no benefit for practitioner adherence	Medium study limitations, precise results

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
					Difference in out-of-home placements and child behavior problem scores at 18 months between ARC-only and usual-care groups did not meet statistical significance (p=0.05).	Low for no benefit for patient change in mental health symptoms at 18 months	Medium study limitations, precise results (small sample size), CIs likely overlap
Program to improve organizational climate and culture Glisson et al., 2012[40,49]	Cluster RCT 352 caregivers of youth ages 5–18 in 18 programs	General mental health problems Ages 8–24 years	Program to improve organizational climate and culture vs. usual care	Educational meetings or materials Educational outreach visits Provider satisfaction initiative Audit and feedback	Trends toward improvement in all domains; nonoverlapping CI for some domains showing significant improvements (p<0.05) for ARC group vs. usual care	Low for benefit for practitioner satisfaction	Medium study limitations, imprecise results (small study sample)
					Lower problem behavior scores for youth in the ARC group compared with those in the control group during first 6 months of followup (following 18-month organizational implementation), effect size=0.29	Low for benefit for patient change in mental health symptoms	Medium study limitations, imprecise results (small study sample)
Training nurses to educate parents about EBPs Gully et al., 2008[44]	Interrupted time series in Study 1, 172 parents or caregivers; RCT in Study 2, 51 parents or caregivers	General mental health symptoms (children suspected of abuse during forensic medical examinations) Ages 2–17 years	Protocol to train nurses to educate parents about EBPs vs. typical services	Educational meetings or materials Educational outreach visits Patient-reported data	Strategy improved parent ratings of access to care (mean difference between groups ranged from 0.08 to 2.1 points in Study 1 and 0.6 to 1.9 in Study 2) (scale=1–5)	Low for benefit for patient access to care	High risk of bias, consistent, direct, precise results

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
					Improved parent ratings of satisfaction of care by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Low for benefit for patient satisfaction	High risk of bias, consistent, direct, precise results
					Improved parent ratings of treatment engagement by a mean of 0.9 in Study 1 and 2.5 in Study 2 (scale=1–5)	Low for benefit for treatment engagement	High risk of bias, consistent, direct, precise results
					Improved parent ratings of therapeutic alliance by a mean of 0.4 in Study 1 and 0.9 in Study 2 (scale=1–5)	Low for benefit for therapeutic alliance	High risk of bias, consistent, direct, precise results
Adding intensive quality assurance to implement an EBP Henggeler et al., 2008[43]	Controlled clinical trial, 30 practitioners, N of caregiver and patient reports and monthly data points NR	Substance use disorders (adolescents with marijuana abuse) Ages 12–17 years	Intensive Quality Assurance (IQA) system vs. workshop only to implement an EBP intervention	Quality monitoring	Study does not provide sufficient detail to judge magnitude of effect on practitioner adherence to cognitive behavioral therapy and monitoring techniques	Insufficient for practitioner adherence and fidelity	High study limitations, imprecise results
Adding computer-assisted training with or without ongoing supervision and coaching to practitioners implementing an EBP Henggeler et al., 2013[48]	Cluster RCT; 161 therapists	Substance use disorders Ages 12–17 years	Workshop and resources (WSR) vs. WSR and computer-assisted training (WSR+CAT) to implement an EBP intervention	Educational meetings or materials	No statistically significant difference between groups for use, knowledge, and adherence	Insufficient for additional benefit of WSR+CAT vs. WSR comparison group for practitioner use, knowledge, and adherence	Medium study limitations, imprecise, small sample sizes, cannot determine whether CIs cross line of no difference

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
			WSR vs. WSR+CAT and supervisory support (WSR+CAT+SS) to implement an EBP intervention	Educational meetings or materials Educational outreach visits	No statistically significant difference between groups for use, knowledge, and adherence	Insufficient for additional benefit of WSR+CAT+SS vs. WSR comparison group on practitioner use, knowledge, and adherence competence/ skills	Medium study limitations, imprecise, small sample sizes, cannot determine if CIs cross line of no difference
Training practitioners to identify and refer cases Lester et al., 2009[37]	Cluster RCT; 110 practices, 179 patients	Psychosis (adolescents and adults with first-episode psychosis) Ages 14–30 years	Professional training to identify and refer cases vs. usual care	Educational meetings or materials Educational outreach visits	Relative risk (RR) of referral to early intervention after first contact: 1.20, 95% CI, 0.74 to 1.95, p=0.48	Insufficient for patient access to care	High study limitations, imprecise results
					No statistically significant differences between groups in changes in patient mental health status	Insufficient for patient change in mental health symptoms	High study limitations, imprecise results
					Patients in the professional training group averaged 223.8 fewer days for time from the first decision to seek care to the point of referral to an early intervention service than patients in the control group	Low for benefit for service utilization	High study limitations, imprecise results
					No adverse events were reported, no significant between-group differences for false-positive referral rates from primary care	Insufficient for patient harms	High study limitations, unknown precision

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Training practitioners with or without feedback to implement an EBP Lochman et al.,2009[46]	Cluster RCT, 511 patients	Externalizing behaviors (children at risk for aggressive behaviors) Ages: third-grade students	Professional training plus feedback (CP-TF) to implement an EBP intervention vs. control	Educational meetings or materials Audit and feedback	Students in CP-TF group had fewer behavioral problems as rated by teachers (beta=-0.41, SE=0.16, p=0.01) than controls but no significant difference in teacher ratings or parent ratings	Low for no benefit for changes in mental health status	Medium study limitations, precise results
					Students in CP-TF group had fewer minor assaults (e.g., hitting or threatening to hit a parent, school staff, or student) as reported by the child (beta=-0.25, SE=0.12, p=0.03) and social/academic competence as reported by the teacher (beta=0.35, SE=0.13, p=0.01) compared with controls	Low for benefit for change in socialization skills and behaviors	Medium study limitations, precise results

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
			Professional training only to implement an EBP intervention (CF-BT) vs. control	Educational meetings or materials	No significant difference in behavioral problems as rated by teachers or parents or student-reported assaults between CP-BT and control groups	Low for no benefit for changes in mental health status	Medium study limitations, precise results
					No significant differences in social/academic competence as reported by the teacher, nor were any significant differences found between groups on social skills as rated by parents.	Low for no benefit for change in socialization skills and behaviors	Medium study limitations, precise results
Training practitioners to use a patient medication monitoring program Ronsley et al., 2012[38]	Interrupted time series Health care practitioners for 2,376 patients	Psychosis Ages <19 years (mean age = 11)	Patient medication monitoring training program for practitioners vs. usual care	Educational meetings or materials Educational outreach visits Reminders	38.3% of patients had a metabolic monitoring and documentation tool (MMT) in the charts after program implementation; drop in the prevalence of second-generation antipsychotic prescribing from 15.4% in the pre-metabolic monitoring training program (MMTP) period to 6.4% in the post-MMTP period ($p<0.001$)	Low for benefit for practitioner adherence	High study limitations, precise outcomes
					Increased metabolic monitoring over time (level of change varied by type of monitoring)	Low for benefit for patient service utilization	High study limitations, precise outcomes

Table C. Strategies to improve mental health of children and adolescents: Summary table (continued)

Strategy, Study	Designs, N	Target Condition and Ages of Youth	Comparisons	Component of the Strategy	Major Findings	Strength of Evidence From Results	Reasons for Strength of Evidence
Embedding a behavioral health care practitioner in primary care Sterling et al., 2015[47]	Cluster RCT, 47 pediatricians with 1,871 eligible patients	Varied conditions among children attending a pediatric primary care office Ages 12–18	Pediatrician only vs. embedded behavioral health care practitioner (BHCP)	Multidisciplinary teams	No significant differences in substance use assessment between study arms (aOR, 0.93; 95% CI, 0.72 to 1.21); patients in the embedded BHCP group more likely than those in the pediatrician-only group to receive brief intervention (aOR=1.74, 95%CI, 1.31 to 2.31); patients in the BHCP group less likely to receive a referral to a specialist than patients in the primary-care-[b] only group (aOR=0.58, 95%CI, 0.43 to 0.78)	Low for no benefit for practitioner adherence (2 of 3 adherence outcomes were statistically significant)	Medium study limitations, unable to assess precision
Co-locating an EBP program in primary care Wildman et al., 2009[41]	Controlled clinical trial, 4 pediatric practices, 20,917 children with primary care visit	Externalizing behavior problems Ages 2–12 years	Colocation of a behavioral health EBP parenting program in primary care vs. enhanced referral to a behavioral health EBP parenting program in a location external to the practice.	Changing the scope of benefits	OR for attending first EBP visit, 3.10; 95% CI, 1.63 to 5.89	Low for benefit for patient access to care	High study limitations, precise results
					No improvement in mean number of sessions attended (calculated mean difference: -1.01; 95% CI, 2.60 to 0.58)	Insufficient for patient service utilization	High study limitations, precise results

[a] Four study groups were examined: ARC+MST, ARC only, MST only, and usual care. Comparisons were ARC only vs. usual care or any ARC (combined ARC+MST and ARC only) vs. no ARC (combined MST and usual care), as noted.

[b] Fewer referrals seen as improvement because this outcome indicates that the practitioner was able to give brief intervention without referral to behavioral health specialists.

ADHD = attention deficit hyperactivity disorder; aOR = adjusted odds ratio; ARC = Availability, Responsiveness and Continuity; CBT = cognitive behavioral therapy; CI = confidence interval; CP-TF = Coping Power training plus feedback; EBP = evidence-based practice, EHR = electronic health record; IAU = implementation as usual; IQA = Intensive Quality Assurance; MMT = metabolic monitoring program; MMTP = metabolic monitoring training program; MST = multisystemic therapy; N = number; NR = not reported; NS = not significant; OR = odds ratio; p = probability; RCT = randomized controlled trial; RR = relative risk; P4P = pay for performance; SE = standard error; WSR = workshop plus resources; WSR+CAT = workshop plus resources plus computer-assisted training; WSR+CAT+SS = workshop plus resources plus computer-assisted training plus supervisory support.

The strongest evidence in the review comes from a study of pay for performance. Therapists in the pay-for-performance group were more than twice as likely to demonstrate implementation competence as were the implementation-as-usual therapists (*moderate strength of evidence of benefit*).[42] Other outcomes for which we found evidence of benefit (*low strength of evidence of benefit*) included:

1. Improved practitioner adherence to EBPs or guidelines from training practitioners to monitor metabolic markers,[38] providing computer decision support plus EHR that included diagnosis and treatment guidelines,[35] and offering an Internet portal for practitioner access to practice guidelines;[45]
2. Improved practitioner morale, engagement, and stress from a program to improve organizational climate and culture;[40]
3. Improved patient access to care, parent satisfaction, treatment engagement, and therapeutic alliance from training nurses to educate parents about EBPs;[44]
4. Improved patient functional status from weekly feedback on patient symptoms and functioning to practitioners;[13] and
5. Improved service utilization from training practitioners about monitoring medications[38] and appropriately identifying and referring patients.[37]

Only four strategies (1 one study each) consistently provided *insufficient or evidence of no benefit* across all reported outcomes. These included:

1. A strategy testing augmented active learning versus computerized routine learning versus routine practitioner workshop to implement an EBP,[39]
2. A collaborative consultation treatment service to promote the use of titration trials and periodic monitoring during medication management versus control,[36]
3. An Intensive Quality Assurance system versus workshop to implement an EBP intervention,[43] and
4. Use of additional computerized assisted training or computerized training plus supervisory support to implement an EBP versus using a workshop and resources only.[48]

The studies varied with respect to the numbers and types of active components; i.e., we observed considerable differences in components in treatment group strategies and comparison group strategies. In some studies, the treatment group contained several components and the comparison group contained none of those components. In other studies, both the treatment and comparison groups tested strategies with multiple components, with varying numbers of differences in components across arms. Because both arms often received active interventions, the Hawthorne effect may explain lack of effectiveness. We did not find any consistent patterns of effectiveness involving the number of active components. That is, we did not find that studies that employed strategies with a single active component had any better or any worse effect on outcomes than those that employed multiple active components.

Additional heterogeneity arose from several other sources and precluded any quantitative synthesis of our findings. Except for two studies reported in one publication[44] and two trials (three publications) reporting variants of a similar intervention,[14,40,49] none of the other studies tested similar strategies. The outcomes of the studies varied widely. Similarly, settings differed greatly (community-based hospitals and clinics, general practice and primary care, home-based mental health systems, schools). Finally, the targets of each strategy, such as practitioners, practices, or systems, also differed considerably.

The absence of evidence on several factors of interest further limited our conclusions. We found no evidence of studies examining several intermediate outcomes, particularly system-level intermediate outcomes. We also identified no studies that measured final patient health outcomes such as co-occurring conditions or mortality. We also found no evidence of strategies testing several components of the EPOC taxonomy, including any regulatory components, and little evidence on strategies with financial components.

Of the 17 studies in our review, one study had low risk of bias and three had medium risk of bias. We rated seven as having unclear risk of bias and six as having high risk of bias. Various issues with study design, attrition, and incomplete information reported by study authors precluded most of these studies from having a low or medium risk of bias.

The uncertain or high risk of bias of most of these studies affected the overall strength of evidence grades, as did the fact that we mainly had only single studies for each strategy examined.

Key Question 2. Harms Associated With Strategies to Improve Mental Health Care for Children and Adolescents

Only one study evaluated the harms associated with professional training to identify and refer cases to early-intervention services for untreated first-episode cases of psychosis.[37] The study reported no adverse events and no differences in false-positive referral rates. We graded the evidence on harms as having insufficient strength, based on high study limitations and imprecise results.

Key Question 3. Moderators of the Effectiveness of Strategies to Improve Mental Health Care for Children and Adolescents

Overall, we found evidence on four strategies that examined moderators of the effectiveness of strategies to improve mental health care for children and adolescents. Three examined whether training intensity influenced the degree of effectiveness; of these, two strategies were graded as having insufficient strength of evidence. The third strategy had low strength of evidence for benefit for patient intermediate outcomes (access to care) and patient health and service utilization outcomes (change in mental health status).

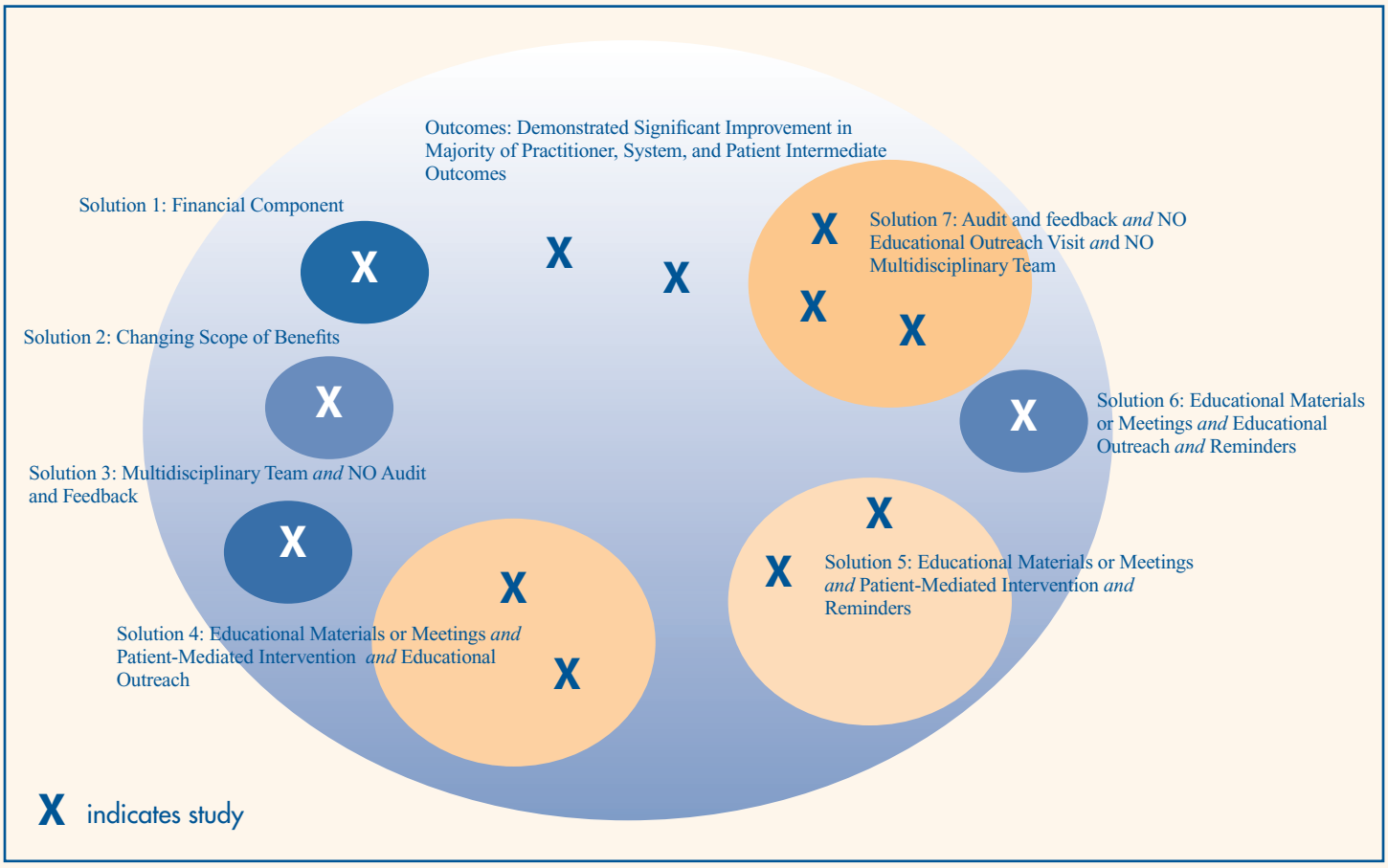
A fourth study examined the moderating effects of fidelity to the EBP (meeting target Adolescent Community Reinforcement Approach) used as part of the strategy. We graded the evidence on the moderating effect of fidelity on this strategy as having low strength for no benefit on patient health outcomes and patient remission status.

We did not find studies that examined most of our previously-specified list of moderators such as patient characteristics, intervention characteristics other than training intensity, factors of the outer or inner setting/organizational factors, characteristics of involved individuals, process characteristics other than fidelity to the training, or other moderators such as length of followup.

Finding Solutions for Success

We turned to QCA to understand what combinations of components ("condition sets") might serve as solutions or "recipes" for success. We examined several different models that contained different combinations of intervention components resulting in two different outcomes. We chose the model that best fit our data with the highest level of consistency (proportion of solutions resulting in success or outcome) and coverage (proportion of observations explained by the solutions). Our model included the presence or absence of several professional components (educational materials or meetings, educational outreach, patient-mediated interventions, audit and feedback), any financial component, organizational structural-oriented components (quality monitoring, change in scope and nature of benefits and services and patient choice of treatment), and organizational provider-oriented component (use of clinical multidisciplinary teams). We defined success as having a statistically significant improvement in either a majority of practitioner-, system-, and patient-level intermediate outcomes or at least one patient health or service utilization outcome showing at least low strength of evidence for benefit. The QCA yielded seven solutions associated with success, described below and shown in Figure C. Four of the solutions included only one study each. Two solutions included two studies each. And one solution included three studies. Two of the studies that showed benefit did not belong to any of the solutions yielded by the QCA. Of note, one study met criteria for two different solutions associated with success.

Figure C. Venn diagram of QCA findings



QCA = Qualitative Comparative Analysis

Our analysis included 17 studies; 12 showed significant improvements (i.e., significant improvement in majority of practitioner, system, or patient intermediate outcomes or at least one patient health or service utilization outcome showing at least low strength of evidence for benefit coded as 1). Five did not.

In the Boolean analysis of the truth table, no conditions were individually necessary or sufficient, and no necessary combinations occurred. Analysis of sufficient combinations for achieving significant improvements showed seven solutions, each with 100-percent consistency. Notably, the model had 83-percent coverage, accounting for 10 of the 12 studies that demonstrated at least low strength of evidence of benefit for at least one outcome. These solutions were:

- Having any financial component; or
- Having a component that included changing the scope or nature of benefits or services and patient choice of treatment; or
- Using clinical multidisciplinary teams and not having an audit and feedback component
- Having educational materials or meetings, patient-mediated interventions, and educational outreach; or
- Having educational materials or meetings, patient-mediated interventions, and reminders; or
- Having educational materials or meetings, educational outreach, and reminders; or
- Having an audit and feedback component and not having educational outreach and not using a clinical multidisciplinary team.

Discussion

Key Findings and Strength of Evidence

Overall, 12 of the 17 studies demonstrated effectiveness as measured by low or moderate strength of evidence for benefit for at least one outcome of interest. Our confidence in these results is limited by the paucity of studies on any strategy. We found moderate strength of evidence of benefit for pay for performance.[42] We found low strength of evidence of benefit for at least one outcome among strategies that contained:

- reminders (i.e., a component that included patient- or encounter-specific information, provided verbally, on paper, or on a computer screen, that was designed or intended to prompt a health professional to recall information),[35,38,45]
- a patient-mediated component (i.e., one that collected new clinical information directly from patients then given to the provider to review),[35,44,45]
- enhanced referrals and patient choice of treatment.[41]
- We found low strength of evidence of no benefit for intermediate outcomes for strategies that included the following combinations of professional components:
- educational materials and/or educational meeting components only[39,48]
- educational materials and outreach components only.[37,46]

We were unable to judge the potential for harms associated with these strategies that may mitigate benefits based on the single included study on early intervention for first-episode psychosis that reported no adverse events and no differences in false-positive referral rates. In addition, the available evidence from four studies on two moderators does not permit us to make general conclusions about the conditions under which these strategies might work optimally.

Applicability

The applicability of findings is limited to professionally trained practitioners of children and adolescents with mental health and/or substance use disorders who delivered QI, implementation, and dissemination strategies in typical service settings. All strategies reviewed were focused at the practitioner (e.g., training practitioners) or system (e.g., implementing a new medical management system) level. Comparison conditions included usual treatment, lower-intensity versions of the strategy under study, and prestrategy implementation cases in one study implementing a system-level strategy within a hospital.

Outcomes examined in the studies included intermediate practitioner, intermediate patient, and a single intermediate system outcomes (uptake). No studies examined other intermediate system outcomes such as feasibility, timeliness, penetration, sustainability, and resources, including costs. Several patient health outcomes of interest such as comorbidity and mortality were not examined in any included studies. Thus, applicability of findings is limited to these outcomes examined.

Limitations of the Systematic Review Process

Challenges in this systematic review arose from the sparse amount of prior literature on this topic that limited defining many of the details of our review a priori. Specifically, we struggled with defining the intervention of interest, constructing the search strategy, and applying prespecified inclusion/exclusion criteria. The lack of consistency in the terminology used in the published literature meant that the use of self-selected descriptors such as "QI," "implementation," or "dissemination" by study authors did not conform to our a priori definitions of these types of studies or to the other similarly labeled studies in the field; this lack of consistency led to our reliance on the EPOC taxonomy as our primary analytic framework. Regarding searches, we ran multiple iterations over a period of 7 months. We initially mirrored the search strategy in a previously published review but had to make substantial changes to capture concepts or terms that were not indexed by the National Library of Medicine's MeSH.

We found that attempts to specify the population and comparison criteria to ensure greater homogeneity of included interventions resulted in additional challenges. For example, our focus on children and adolescents with existing mental health issues (rather than the risk of mental health issues only) did not enable focus on prevention. In addition, although we included a broad range of eligible comparators in our protocol (usual care, or any other QI, implementation, or dissemination strategy), we did encounter otherwise eligible studies in which the intervention combined both a patient-level intervention and a system-level strategy to implement or disseminate that intervention. Because the use of a usual-care arm did not permit the authors to draw conclusions about the effect of the implementation or dissemination strategy apart from the underlying intervention, we excluded these studies for having a wrong comparator.[50-57]

Limitations of the Evidence Base

We found relatively few studies that examined the effectiveness of strategies to improve the mental health care of children and adolescents, Although we did find evidence that some strategies are effective in improving both intermediate and patient health and resource utilization outcomes, we found only one study that focused on system-level intermediate outcomes and none that compared the costs of these strategies.

The lack of a common language to describe even a basic concern such as the primary purpose of the strategies (QI, implementation, or dissemination) served as a hindrance to synthesis. Strategies varied significantly in the number of components; the reporting on these components was not always clear enough to adequately describe the strategy or fully understand the relative importance of component parts. Studies often offered limited descriptions of "usual-care" arms when compared with descriptions of experimental arms. Even with limited reporting, we found wide differences in the number, intensity, and services offered in "usual-care" arms. These differences sharply limited our ability to make statements about the overall effectiveness of these strategies as a class.

Only one study examined harms. Although the field generally acknowledges the vast array of potentially influential moderators in implementation research,[58] we uncovered only four studies on two moderators (intensity and fidelity). The paucity of evidence on these issues further limits our understanding of the minimum change in strategy needed to achieve a significantly different process or health outcome.

We rated most outcomes as insufficient or low strength of evidence because of the underlying heterogeneity or limited number of studies on specific strategy types, system or practitioner targets, or child or adolescent conditions. In some instances, our grades were limited by high risk of bias in included.

Our ability to derive firm conclusions on the effectiveness of included strategies was also hindered by reporting issues in the literature. Authors reported complex analyses but often did not report other issues well enough to permit an independent evaluation of the effect size,[46] precision of the effect,[35-37,40] or risk of bias.[35,46]

Research Recommendations

The evidence base is marked by a small number of studies on diverse strategies and outcomes focusing on intermediate and health outcomes and resource use; we had very few studies on harms or moderators. Our review highlights the fact that the current state of the evidence does not give clinicians and health plan administrators a definitive understanding of best methods to introduce EBPs successfully into clinical settings. Third-party payers are paying increasing attention to quality metrics, as health care systems move to accountable care models. We found no studies on regulatory components and just one study testing the effectiveness of a financial component, specifically for pay for performance.[45] Future research efforts should evaluate variations of such programs according to patient, provider, organization, systems, and setting characteristics. A better understanding of these variables can impede or promote the implementation and dissemination of EBPs.

We did not find evidence on the majority of the outcomes that we specified a priori. Of particular note, seven strategies (two from a single publication) relied on EBPs; for that reason, these investigators did not report patient health outcomes.[39,43-45,47,48] When researchers maintain fidelity to the original intervention, the assumption that the same level of effectiveness will occur in a new trial is reasonable and leads to an efficient use of research funds. Unfortunately, not all studies measured fidelity adequately. New strategies relying on EBPs must, at a minimum, report on fidelity so practitioners and policymakers can judge whether the strategy is, in fact, new intervention, rather than implementation or dissemination of an existing intervention. Information on pragmatic issues related to implementation (fidelity, adaptation, and minimum elements necessary to achieve change) may not necessarily require new studies on strategies with existing information; support of analyses done with data from existing studies may fill some of the gap.

The majority of included studies appropriately used cluster RCTs. Cluster RCTs, like pragmatic trials, need more resources than conventional RCTs and are harder to complete than conventional studies. An additional consideration of cluster RCTs relates to reporting. The studies we found were marked by poor reporting or failure to report key details of the strategy or differences across study arms. Concerns about the inadequacies of reporting have been noted elsewhere in the literature.[59,60] A recent tool, the StaRI, (standards for reporting implementation studies of complex interventions), offers standards for reporting implementation studies that, if adopted widely, can significantly improve the utility of these studies and the pace of translation of evidence into practice.[61]

Although the failure to use EBPs results can lead to gaps between potential and achieved outcomes, closing such gaps requires more than just using an array of EBPs.

What continues to be unknown is how to bridge the gap in the context of the finite resource of time allocated for a patient encounter. As expectations for documenting or checking off quality metrics for each action within a patient encounter increase, the risk of errors of omission or commission increases. For new information to be actionable, more evidence is needed on the relative merits of each action or strategy.

More research is needed on strategies for the QI, implementation, and dissemination of EBPs in psychotherapy treatments as well as medication treatments of mental illness in youth. Other important targets include the development of dissemination strategies for introducing mental health care into areas lacking in mental health care, for example, very rural areas with fewer mental health providers. In these areas especially, targeting primary care providers may be essential.

Conclusions

Our findings suggest that several approaches can improve both intermediate and final health outcomes and resource use. Twelve of the 17 included studies (11 of the 16 strategies) significantly improved at least one such outcome or measure. Moderate strength of evidence (from one RCT) supported using provider financial incentives such as pay-for-performance to improve the competence with which practitioners can implement EBPs. We found inconsistent evidence involving strategies with educational meetings, materials, and outreach; programs appeared to be successful in combination with reminders or providing practitioners with newly collected clinical information. We also found low strength of evidence for no benefit for initiatives that included only educational materials or meetings (or both), or only educational materials and outreach components.

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. PMID: 23677130.
3. 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. PMID: 7777649.
4. Dishion TJ, McCord J, Poulin F. When interventions harm. Peer groups and problem behavior. Am Psychol. 1999 Sep;54(9): 755-64. PMID: 10510665.
5. 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.
6. 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.
7. 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. PMID: 11533391.
8. 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. PMID: 22025589.
9. 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. PMID: 12202276.
10. 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. PMID: 24399817.
11. 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.
12. 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.
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. 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. 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. Institute of Medicine. Crossing the Quality Chasm. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.

18. 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. PMID: 19104929.

19. 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. PMID: 22726759.

20. 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. PMID: 15224451.

21. 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. PMID: 18085434.

22. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy. Oslo: Norwegian Knowledge Centre for the Health Services; 2002 https://epoc.cochrane.org/epoc-taxonomy.

23. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339(4):b2535. PMID: 19622551.

24. 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/

25. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928. PMID: 22008217.

26. 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.

27. Sterne JAC, Higgins JPT, Reeves BC, et al. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI), Version 1.0.0. 2014 24 September 2014. http://www.riskofbias.info. Accessed 29 May 2015.

28. 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/ PMID: 21433337.

29. Booth A, Harris J, Croot E, et al. Towards a methodology for cluster searching to provide conceptual and contextual "richness" for systematic reviews of complex interventions: case study (CLUSTER). BMC Med Res Methodol. 2013;13:118. PMID: 24073615.

30. 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

31. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2014 Dec 20. PMID: 25721570.

32. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011 Dec;64(12):1283-93. PMID: 21839614.

33. Bhandari M, Sprague S, Hanson B, et al. Health-related quality of life following operative treatment of unstable ankle fractures: a prospective observational study. J Orthop Trauma. 2004 Jul;18(6):338-45. PMID: 15213498.

34. 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/

35. Carroll AE, Bauer NS, Dugan TM, et al. Use of a computerized decision aid for ADHD diagnosis: a randomized controlled trial. Pediatrics. 2013 Sep;132(3):e623-9. PMID: 23958768.

36. Epstein JN, Rabiner D, Johnson DE, et al. Improving attention-deficit/hyperactivity disorder treatment outcomes through use of a collaborative consultation treatment service by community-based pediatricians: a cluster randomized trial. Arch Pediatr Adolesc Med. 2007 Sep;161(9):835-40. PMID: 17768282.

37. Lester H, Birchwood M, Freemantle N, et al. REDIRECT: cluster randomised controlled trial of GP training in first-episode psychosis. Br J Gen Pract. 2009 Jun;59(563):e183-90. PMID: 19520016.

38. Ronsley R, Rayter M, Smith D, et al. Metabolic monitoring training program implementation in the community setting was associated with improved monitoring in second-generation antipsychotic-treated children. Can J Psychiatry. 2012 May;57(5):292-9. PMID: 22546061.

39. Beidas RS, Edmunds JM, Marcus SC, et al. Training and consultation to promote implementation of an empirically supported treatment: a randomized trial. Psychiatr Serv. 2012;63(7):660-5. PMID: 22549401.

40. Glisson C, Hemmelgarn A, Green P, et al. Randomized trial of the Availability, Responsiveness, and Continuity (ARC) organizational intervention with community-based mental health programs and clinicians serving youth. J Am Acad Child Adolesc Psychiatry. 2012 Aug;51(8):780-7. PMID: 22840549.

41. Wildman BG, Langkamp DL. Impact of location and availability of behavioral health services for children. J Clin Psychol Med Settings. 2012 Dec;19(4):393-400. PMID: 23053830.

42. Garner BR, Godley SH, Dennis ML, et al. Using pay for performance to improve treatment implementation for adolescent substance use disorders: results from a cluster randomized trial. Arch Pediatr Adolesc Med. 2012 Oct;166(10):938-44. PMID: 22893231.

43. Henggeler SW, Sheidow AJ, Cunningham PB, et al. Promoting the implementation of an evidence-based intervention for adolescent marijuana abuse in community settings: testing the use of intensive quality assurance. J Clin Child Adolesc Psychol. 2008 Jul;37(3):682-9. PMID: 18645758.

44. Gully KJ, Price BL, Johnson MK. Increasing abused children's access to evidence-based treatment: diffusion via parents as consumers. Child Maltreat. 2008 Aug;13(3):280-8. PMID: 18359928.

45. Epstein JN, Langberg JM, Lichtenstein PK, et al. Use of an Internet portal to improve community-based pediatric ADHD care: a cluster randomized trial. Pediatrics. 2011 Nov;128(5):e1201-8. PMID: 22007005.

46. Lochman JE, Boxmeyer C, Powell N, et al. Dissemination of the Coping Power program: importance of intensity of counselor training. J Consult Clin Psychol. 2009 Jun;77(3):397-409. PMID: 19485582.

47. Sterling S, Kline-Simon AH, Satre DD, et al. Implementation of screening, brief intervention, and referral to treatment for adolescents in pediatric primary care a cluster randomized trial. JAMA Pediatrics. 2015;169(11):e153145. PMID: CN-01104952.

48. Henggeler SW, Chapman JE, Rowland MD, et al. Evaluating training methods for transporting contingency management to therapists. J Subst Abuse Treat. 2013 Nov-Dec;45(5):466-74. PMID: 23910392.

49. Glisson C, Hemmelgarn A, Green P, et al. Randomized trial of the Availability, Responsiveness and Continuity (ARC) organizational intervention for improving youth outcomes in community mental health programs. J Am Acad Child Adolesc Psychiatry. 2013 May;52(5):493-500. PMID: 23622850.

50. Asarnow JR, Jaycox LH, Duan N, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. JAMA. 2005 Jan 19;293(3):311-9. PMID: 15657324.

51. Asarnow JR, Jaycox LH, Tang L, et al. Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. Am J Psychiatry. 2009 Sep;166(9):1002-10. PMID: 19651711.

52. Ngo VK, Asarnow JR, Lange J, et al. Outcomes for youths from racial-ethnic minority groups in a quality improvement intervention for depression treatment. Psychiatr Serv. 2009 Oct;60(10):1357-64. PMID: 19797376.

53. Wells KB, Tang L, Carlson GA, et al. Treatment of youth depression in primary care under usual practice conditions: observational findings from Youth Partners in Care. J Child Adolesc Psychopharmacol. 2012 Feb;22(1):80-90. PMID: 22251025.

54. Henggeler SW, McCart MR, Cunningham PB, et al. Enhancing the effectiveness of juvenile drug courts by integrating evidence-based practices. J Consult Clin Psychol. 2012 Apr;80(2):264-75. PMID: 22309470.

55. Liddle HA, Dakof GA, Henderson C, et al. Implementation outcomes of Multidimensional Family Therapy-Detention to Community: a reintegration program for drug-using juvenile detainees. Int J Offender Ther Comp Criminol. 2011 Jun;55(4):587-604. PMID: 20427547.

56. Weisz JR, Southam-Gerow MA, Gordis EB, et al. Cognitive-behavioral therapy versus usual clinical care for youth depression: an initial test of transportability to community clinics and clinicians. J Consult Clin Psychol. 2009 Jun;77(3):383-96. PMID: 19485581.

57. Bryant J, Passey ME, Hall AE, et al. A systematic review of the quality of reporting in published smoking cessation trials for pregnant women: an explanation for the evidence-practice gap? Implement Sci. 2014;9:94. PMID: 25138616.

58. 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. PMID: 19664226.

59. Rycroft-Malone J, Burton CR. Is it time for standards for reporting on research about implementation? Worldviews Evid Based Nurs. 2011;8:189-90.

60. Newhouse R, Bobay K, Dykes PC, et al. Methodology issues in implementation science. Med Care Res Rev. 2013;51 (Suppl 2):S32-40.

61. Pinnock H, Epiphaniou E, Sheikh A, et al. Developing standards for reporting implementation studies of complex interventions (StaRI): a systematic review and e-Delphi. Implement Sci. 2015;10(1):42. PMID: 25888928.

Full Report

This executive summary is part of the following document: Forman-Hoffman VL, Cook Middleton J, McKeeman JL, Stambaugh LF, Christian RB, Gaynes BN, Kane HL, Kahwati LC, Lohr KN, Viswanathan M. Strategies To Improve Mental Health Care for Children and Adolescents. Comparative Effectiveness Review No. 181. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2012-00008-I.) AHRQ Publication No. 16-EHC035-EF. Rockville, MD: Agency for Healthcare Research and Quality; December 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

AHRQ Pub. No.16-EHC035-EF
December 2016
www.ahrq.gov