Journal of Pediatric Psychology, 2024, **00**, 1–18
https://doi.org/10.1093/jpepsy/jsae038
Systematic Review

A systematic review and meta-analysis of pediatric integrated primary care for the prevention and treatment of physical and behavioral health conditions

Cody A. Hostutler [1,*] PhD, Jeffrey D. Shahidullah,[2] PhD, Jennifer A. Mautone,[3] PhD, Tiffany M. Rybak [4,5] PhD, Chimereodo Okoroji,[4,5] PhD, Teryn Bruni,[6] PhD, Kevin G. Stephenson, PhD, Leah Vance Utset, MD, Kurt A. Freeman,[7] PhD, Leah LaLonde,[1] PhD, Andrew R. Riley [7] PhD

[1]Pediatric Psychology and Neuropsychology, Nationwide Childrens Hospital, The Ohio State University, Columbus, OH, United States
[2]Department of Psychiatry and Behavioral Sciences, Dell Medical School, The University of Texas at Austin, Austin, TX, United States
[3]Department of Child & Adolescent Psychiatry & Behavioral Sciences, Children's Hospital of Philadelphia & Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[4]Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[5]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States
[6]Department of Psychology, Algoma University, Marie, ON, Canada
[7]Institute on Development and Disability, Department of Pediatrics, Oregon Health & Science University, Portland, OR, United States

*Corresponding author: Cody A. Hostutler, Department of Pediatric Psychology and Neuropsychology, Nationwide Children's Hospital and Department of Pediatrics at The Ohio State University, 700 Childrens Drive, Columbus, OH 43205, United States. Email: Cody.Hostutler@nationwidechildrens.org

Abstract

Objective: To evaluate the effects of behavioral health interventions delivered within pediatric integrated primary care models on clinical outcomes.

Methods: We searched Medline, EMBASE, CENTRAL, PsycINFO, and SCOPUS for studies published from January 1, 1998, to September 20, 2023. We included studies that evaluated onsite behavioral health integration in pediatric primary care using a comparator condition (usual, enhanced usual care, or waitlist). Outcome data on symptom change, impairment/quality of life, health indicator, and behavior change were extracted using Covidence software. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guideline was followed Risk of bias analysis was conducted using the Cochrane Risk of Bias tool. We used multilevel meta-analysis to synthesize multiple outcomes nested within studies. Open Science Foundation pre-registration: #10.17605/OSF.IO/WV7XP.

Results: In total, 33 papers representing 27 studies involving 6,879 children and caregivers were included. Twenty-four studies were randomized controlled trials and three were quasi-experimental designs. Seventeen papers reported on treatment trials and 16 reported on prevention trials. We found a small overall effect size (SMD = 0.19, 95% confidence interval [0.11, 0.27]) supporting the superiority of integrated primary care to usual or enhanced usual care. Moderator analyses suggested similar effectiveness between co-located and integrated models and no statistically significant differences were found between treatment and prevention trials.

Conclusions: Results suggest that integrated primary care is superior to usual and enhanced usual care at improving behavior, quality of life, and symptoms. Integrated primary care research needs improved standards for reporting to promote better synthesis and understanding of the literature.

Keywords: primary care, pediatrics, integrated care, collaborative care, behavioral health.

In the United States, approximately 20% of youth experience a behavioral health (BH) condition that impairs functioning (Merikangas et al., 2010), and 40% of school-aged children live with at least one chronic health condition (National Survey of Children's Health, n.d.). Children with and without special health care needs are also exposed to a litany of social stressors (e.g., racism, caregiver incarceration, community violence, food insecurity, bullying) that threaten their long-term functioning and health (Copeland et al., 2018). Despite the high prevalence of psychosocial need among young people, access to quality healthcare services is often limited and inequitable (Bailey et al., 2017; Merikangas et al., 2011; Whitney & Peterson, 2019). A minority of children with an identified BH concern ever receive treatment services (Chisolm et al., 2009; Duong et al., 2021; Hacker et al., 2006; Wang et al., 2023), and children from minoritized groups face disproportionate barriers to accessing care (Hadland et al., 2016; McLaughlin et al., 2010; Trent et al., 2019; Whitney & Peterson, 2019). Recognition of these systemic challenges has led to the promulgation of integrated pediatric primary care as a critical strategy for improving access to needed BH services and reducing health disparities (Bagalman et al., 2022; Shahidullah et al., 2023).

Theoretical model

Integrated primary care (IPC) is theorized to improve population health through multiple active mechanisms of change. Receiving care in a familiar setting may reduce

Received: February 19, 2024. **Revised:** May 8, 2024. **Accepted:** May 9, 2024

BH stigma and be preferable to those with more medicalized views of BH (e.g., Miller-Matero et al., 2019). IPC may also reduce the negative impact of social determinants of health by offering a "one-stop shop" for services where patients can see their primary care and BH clinician in the same location, and in some models, during the same visit (Chakawa et al., 2020; Hodgkinson et al., 2017; Riley et al., 2022). Further, IPC builds the capacity of pediatric primary care to improve the health of a defined population using a population health approach, expanding reach and facilitating the provision of more accessible and equitable services (Hostutler et al., 2023; Shahidullah et al., 2023). IPC also increases the collaboration between primary care and BH clinicians allowing for more informed and comprehensive treatment planning to better target patient needs (Olufs et al., 2016; Stancin & Perrin, 2014).

Existing models of integrated primary care

Although the practice of pediatric IPC is not new (Smith et al., 1967), it has recently proliferated (Richman et al., 2020) following the passage of the Affordable Care Act (Public Law No: 111–148, Mar 23, 2010). Integrated primary care practice occurs across a spectrum of "levels of integration" (Table 1; Heath et al., 2013), which intersect with different categorical models (e.g., Primary Care Behavioral Health [PCBH; Reiter et al., 2018], Collaborative Care [CoCM; Unützer et al., 2013]). Those models are comprised of various components (e.g., team-based visits, warm-handoffs, curbside consultations, standalone BH sessions, use of registries, psychiatric consultation), within which specific interventions (e.g., Cognitive Behavioral Therapy) are delivered.

Previous meta-analytic and systematic reviews

As the practice of IPC has grown, so has the body of research evaluating its effects. Asarnow et al. (2015) completed the seminal pediatric IPC meta-analysis of 31 studies involving over 13,000 participants and found that youth receiving IPC treatment had a 66% probability of better outcomes compared to treatment as usual ($d = 0.32$, 95% [confidence interval], 0.21–0.44). Interventions that focused on treatment (versus prevention)and targeted mental health symptoms (versus substance abuse) produced relatively larger effect sizes. Outcomes were not moderated by age, presenting concern, or level of integration (consultation and co-located versus partially and fully integrated models). More recently, Yonek et al. (2020) attempted to conduct a meta-analytic component analysis of integrated models, but instead completed a systematic review due to an insufficient number of trials. Their systematic review of eleven IPC trials found that population-based care, measurement-based care, and use of evidence-based mental health services were the most common components in trials with positive outcomes. Additional systematic reviews (Burkhart et al, 2020; McLeigh et al., 2022; Shahidullah et al., 2018) support the conclusion that IPC models appear to improve treatment access, engagement, and behavioral health outcomes across a range of conditions. It is important to note, however, that a recent systematic review highlighted significant gaps in the reporting and evaluation of external validity in IPC trials, including a lack of reporting on the representativeness of study participants (Callejo-Black et al., 2020).

Literature gaps

Despite a growing body of evidence, pediatric IPC still faces many unanswered questions. In addition to being nearly a

Table 1. Substance abuse and mental health services administration levels of integration.

Level of integration		Description	Indicators
Coordinated	Level 1	Minimal collaboration	• Separate facilities • Separate systems • Minimal or absent communication • Limited understanding of others' roles
	Level 2	Basic collaboration at a distance	• Separate facilities • Separate systems • Some communications driven by patient needs • Appreciation of others' roles
Co-located	Level 3	Basic collaboration on site	• Same facilities, possibly separate space • Separate systems • Some communication and collaboration • Occasional direct contact
	Level 4	Close collaboration on site with some system integration	• Same facilities, same space • Some shared systems • Frequent communication and collaboration • Regular in-person contact
Integrated	Level 5	Close collaboration approaching an integrated practice	• Same facilities, same space • Working to integrate systems • Frequent in-person communication • Regular collaborative meetings • In-depth understanding of roles and culture
	Level 6	Full collaboration in a transformed/merged integrated practice	• Same facilities, same space • Fully integrated systems • Consistent communication at individual team, and system levels • Formal and informal meeting to support the model of integration • Roles and cultures that blur and blend

decade old, Asarnow et al.'s (2015) seminal meta-analysis defined integrated primary care broadly and included interventions that occurred outside of primary care centers (e.g., school-based mental health centers; Mufson et al., 2004), studies assessing the effectiveness of training primary care clinicians to deliver behavioral health interventions (e.g., Wissow et al., 2008), and computer-based interventions (e.g., Walton et al., 2013, 2014). There has also been considerable growth in the field over the last decade, including changes in the lexicon used to describe IPC models, making it challenging to interpret prior comparisons of different models of integrated primary care. In addition, prior reviews have overlooked the stated potential of IPC to improve physical health and have used traditional meta-analytic methodologies that required averaging across outcome types rather than differentiating effectiveness by outcome type (e.g., symptom change, quality of life, physical health indicators, behavior change).

Objectives

To best reflect the breadth of IPC models that have been studied in clinical trials and goals of the primary care system of care, the aims of this systematic review and multilevel meta-analysis are to determine whether IPC leads to improved health outcomes compared to usual or enhanced usual care for children and adolescents (0–21 years). In addition to obtaining an overall effect across outcome types, this multilevel approach allows us to compare the effectiveness of IPC within specific outcome types (i.e., symptom change, quality of life, behavior change, physical health indicator) that are often nested within studies instead of having to select one outcome or average across them. We also assess whether effectiveness varies by level of integration (i.e., co-located vs. integrated), trial type (i.e., prevention, treatment), treatment target (i.e., mental health intervention, substance use, physical health intervention, development), trial design type (RCT, other), risk of bias (high vs. some/low), or participant demographics (i.e., age, race and ethnicity). We chose to focus on level of integration rather than discrete categorial models, because the SAMSHA levels of integration framework is well-established and the categorical models are difficult to measure and classify due to limited of models reporting within the literature and changes in the lexicon over time

Methods

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement guideline in conducting this meta-analysis (Page et al., 2021). The review protocol was pre-registered in Open Science Framework (#10.17605/OSF.IO/WV7XP). Data collection forms, extracted data (including those used in analyses), analytic code, and other tools used in the review are available from the first author.

Eligibility criteria

Study type

We included peer-reviewed trials that incorporated a comparator condition and were conducted in primary medical care settings, including but not limited to randomized controlled trials. Trials with fewer than 10 participants in the treatment or comparator arms at the conclusion of the study were excluded. Non-peer-reviewed publications, conference abstracts, dissertations, and trials without a comparator condition were excluded to maximize quality and minimize bias. As models of primary care and health payment vary substantially across countries, we excluded studies conducted outside the United States to ensure that findings generated from this review are relevant to clinical practice and policy in U.S. systems of care.

Participants

Though there is variability in definitions of "pediatrics", we included youth from birth to 21 years of age consistent with Bright Futures and Food and Drug Administration definitions (Hardin et al., 2017). We made no inclusion restrictions related to health status or condition. Subsets of eligible participants from studies of both pediatric and adult populations were included only if characteristics and outcomes of those participants who were eligible for this analysis could be extracted separately. Study authors were contacted if eligibility based on participant characteristics was unclear.

Interventions

For the purpose of this review, we defined IPC as requiring the integration of interventionists trained to deliver onsite BH services within primary care practices to pediatric patients and their caregivers. This included models of integration that met levels 3–6 of the SAMSHA Framework for Levels of Integrated Healthcare (Heath et al., 2013). Levels 1 and 2 were excluded because these levels describe *off-site* collaboration between behavioral health providers and primary care providers (see Table 1 for details on all six levels). Models were coded as *co-located* if meeting levels 3 and 4 criteria and *integrated* if meeting levels 5 and 6 criteria (e.g., PCBH, CoCM). Studies that recruited from primary care but only delivered services off-site, trained primary care clinicians to deliver interventions themselves, or provided off-site consultation and support to primary care clinicians were excluded. Studies conducted in school-based health centers and those conducted in controlled research settings (e.g., academic research lab) were excluded. Studies exclusively evaluating integration within specialty care settings were also excluded.

Both pediatric and family medicine primary care practices were included (assuming a pediatric sample could be extracted with a maximum age of 21). Studies including a wide range of prevention and intervention strategies targeting physical health, mental health, and substance use within the context of IPC were included. Comparator interventions included usual care (e.g., primary care without integrated BH providers) or enhanced usual care (e.g., primary care provider education, facilitated referral pathways).

Outcomes

Because the goal of pediatric IPC is to improve child health at the population level, we included both physical health and behavioral health outcomes. We extracted clinical outcomes including: symptom change (e.g., Patient Health Questionnaire-9 [PHQ-9; Richardson et al., 2010]), impairment (e.g., impairment rating scales, "healthy days" estimates), quality of life (e.g., Pediatric Quality of Life Inventory [Varni et al., 2005]), health indicator (e.g., body mass index, Hemoglobin A1c, sleep duration), and behavior change (e.g., increased physical activity, acquisition of parenting skills). We did not extract non-clinical outcomes (e.g., access, utilization, satisfaction). We extracted only the primary outcome(s) that were identified by the study authors. When no outcomes were

identified as primary, we extracted all outcomes and accounted for dependency among study effects using multilevel meta-analysis. We similarly used the primary assessment endpoint as defined by the study; if no endpoint was defined as primary, we selected the first outcome assessment post-intervention completion.

Search methods

We searched the following electronic databases for studies published over the last 25 years (from January 1, 1998, to September 20, 2023): Medline, EMBASE, CENTRAL, PsycINFO, and SCOPUS. We reviewed prior systematic reviews and meta-analyses to confirm that this date range would not omit any relevant studies (Asarnow et al., 2015; Callejo-Black et al., 2020; McLeigh et al., 2022; Yonek et al., 2020). We used a comparable search strategy across databases (see Supplementary Appendix A for complete search strategy). We also reviewed the lists of references from included studies (i.e., ancestral search) and the prior systematic reviews and meta-analyses referenced above to identify additional studies that may have been missed in database searches.

Selection process

Studies were imported into Covidence Software (2019) and at least two authors independently screened all article titles and abstracts for inclusion and exclusion criteria and classified them as "yes," "no," or "maybe" for inclusion. Abstracts identified as "yes" and "maybe" underwent full-text review by at least two reviewers. A third author adjudicated discrepancies and consulted the full authorship team for consensus if uncertain.

Data extraction and management

The following data were extracted: authorship, title, year, funding sources, author declarations of interest, sample size, age, sex, SES/insurance status (e.g., Medicaid), race and ethnicity, level of integration, integration model (when possible), intervention (e.g., CBT, parent management training), received dose and duration, delivery mode (e.g., individual, group, telehealth), interventionist type (e.g., care manager, psychologist, trainee), type of comparison and content, study design, randomization method, assessment points, and the outcomes described in the outcomes section. Two authors independently extracted data using an extraction form created within Covidence. The extracting dyad attempted to resolve any differences to reach a consensus; if a consensus could not be reached, the lead author adjudicated the disagreement. If data were unclear or lacking information required for analysis, we sent at least one email to the corresponding author of the study and asked that they respond within three weeks to provide clarity. If we did not hear back from the corresponding author, we labeled the variable as missing. When there were multiple reports of the same study, we amalgamated into a single study summary, treating multiple outcomes as nested within the large study group.

Risk of bias assessment

Study risk of bias

We assessed risk of bias using the Cochrane Risk of Bias version 2 for individual (RoB 2) and cluster randomized trials (RoB 2 CRT; Sterne et al., 2019). The RoB 2 and RoB 2 CRT evaluate the risk of bias in the randomization process, deviations from intended intervention, missing outcome data, measurement of the outcome, and selection of the reported result and classifies the overall risk of bias as having low, high, or unclear risk of bias. We used the ROBINS-I (Sterne et al., 2016) for non-randomized trials. The ROBINS-I rates the risk of bias in overall confounding, selection of participants into the study, classification of interventions, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result to determine an overall risk of bias level (low, moderate, serious, and critical risk of bias). Two authors independently assessed included papers for risk of bias and discrepancies were adjudicated through discussion with at least one other author. We assessed at the paper level rather than the study level as individual papers included different measurement and selection of outcomes that have different risk of biases associated with them.

Reporting bias

To evaluate the possibility of reporting bias (e.g., lack of unpublished results included in the meta-analysis), we employed visual inspection of a funnel plot of the observed outcomes and standard errors and also applied Egger's Test adapted for multilevel meta-analysis according to the recommendations of Rodgers and Pustejovsky (2021) which involves fitting a multilevel meta-analysis with SMD variances or the inverse sample size as moderators. Given the continuous nature of the outcomes, we used the inverse sample size as the chosen predictor for Egger's test (Doleman et al., 2020).

Data analytic approach

We calculated the standardized mean difference (SMD), specifically Hedges' g, for primary endpoint differences between intervention and comparison groups using raw means and standard deviations when available. If other reported information was available which allowed for a calculation of an SMD (e.g., dichotomous outcomes, t-test p-values and sample sizes, means with confidence intervals), we did so using a publicly available online calculator (Wilson, 2023). We conducted a multilevel (mixed effects) meta-analysis to account for non-independence of effect sizes due to using multiple effect sizes from the same study (e.g., including multiple outcome types within the same analysis, nested within each study). We chose this approach because the traditional approach of selecting only one outcome limits the conclusions that can be drawn from the literature, and treating multiple outcomes within studies as separate effect sizes violates the assumption of independence of observations (Cheung, 2019). Random factors included outcomes nested within studies. We fit a correlated and hierarchical effects (CHE) model with Robust Variance Estimation (RVE) and correction for small samples, due to having fewer than 40 studies included (Pustejovsky & Tipton, 2022; Tipton & Pustejovsky, 2015). Within a CHE model, an assumption is that effect sizes within studies (clusters) are correlated and that the correlation is constant both within and across studies. Although this assumption cannot be directly tested, we assumed a moderate—large correlation ($\rho = 0.5$) given the range of outcome domains covered in the meta-analysis. To explore this assumption, we also completed sensitivity analyses (e.g., Pustejovsky & Tipton, 2022) across a range of correlation levels ($\rho = 0.3$–0.7). To help identify potential outliers and influential data points, we calculated Cook's distances for all included outcomes.

Heterogeneity was assessed using the I^2 statistic. For multilevel meta-analysis models, there are multiple I^2 statistics for within- ($I^2_{\text{Level 2}}$) and between-study ($I^2_{\text{Level 3}}$) heterogeneities, and we report the results of both. We tested a number of potential moderators of potential differences in effects sizes between studies. Potential moderators included outcome type (behavioral, health, QoL/impairment, symptom), intervention target (development, mental health, physical health, substance use), intervention type (treatment, prevention), and collaborative model (co-located, integrated). We also evaluated the impact of study design (e.g., randomized vs. non-randomized) and risk of bias (high risk vs. low risk/some risk) on observed effect size. Data did not allow age (overlapping age ranges that could not be cleanly categorized) or race and ethnicity (limited number of studies reporting outcome data by race and ethnicity) to be included within moderator analyses. Meta-analysis was completed using the *metafor* package (Viechtbauer, 2010) in R (v. 4.3.2). Robust confidence intervals for pooled effect sizes were calculated using the *clubSandwich* package in R (Pustejovsky, 2023).

Results

Study selection

Our search strategy yielded 9,188 abstracts, 117 of which received full-text review, and 46 of which met full inclusion criteria, representing 29 different studies. Of the 46 papers that met inclusion criteria, three (Asarnow et al., 2009; Brent et al., 2020; Cates et al., 2018) were follow-up studies that repeated measurement at a longer endpoint and five (Caughy et al., 2003, 2004; Kolko et al., 2020; Ngo et al., 2009; Schwartz et al., 2023) conducted subsample analyses that violated independence of samples assumption in ways that could not be accounted for and were therefore excluded from the meta-analysis. Six studies (Bai et al., 2018; Lavigne et al., 2008, 2010; Mason et al., 2011, 2019; Shaffer et al., 2017) reported incomplete outcome data (e.g., sample size, means, standard deviations), discordant data in different parts of the manuscript (e.g., table versus narrative text), or scores that appeared inaccurate (e.g., scores that were outside the range of the measure without explanation). We emailed the corresponding author at least once requesting clarification or missing data. One author responded with the necessary data in response to our request (Bai et al., 2018). Thus, our meta-analysis included 6,879 participants from 33 papers representing 27 studies (24 RCTs, 3 quasi-experimental). We were only able to extract and compile demographic data from 19 studies (70.4%; $n = 4{,}116$). For these 19 studies, 53.7% of participants were reported as female, 36.5% were non-Hispanic White, 20.1% were Hispanic/Latino, 33.2% Black or African American, 2.4% Asian, 0.9% multiracial, and 6.9% other. See Supplementary Appendix B for study details. Figure 1 displays the full PRISMA flow diagram for study inclusion.

Of the 33 included papers, 23 were categorized as "co-located" and 10 were categorized as "integrated." Limited descriptions of models made categorizing models challenging; however, it appeared that three studies met criteria for CoCM (i.e., care manager, registry, psychiatric consultation; Kolko et al., 2012, 2014; Richardson et al., 2014) and none met criteria for PCBH as described. Most papers ($n = 19$) targeted mental health concerns, four targeted substance use, two targeted physical health, and eight targeted developmental outcomes. We classified 17 papers as treatment trials and 16 as prevention. Comparator conditions included treatment as usual ($n = 16$), enhanced usual care ($n = 15$), and waitlist control ($n = 2$). The most common interventions were CBT ($n = 9$) and parent management training ($n = 7$). Studies used a wide variety of interventionists with eight studies employing a variety of different professionals as interventionists within the same study. Overall study characteristics are summarized in Table 2.

Results of syntheses

The results of individual outcomes within studies are presented in Figure 2 (for treatment studies) and Figure 3 (for prevention studies). The SMDs for individual outcomes ranged from –0.32 to 1.68. The pooled SMD for the three-level meta-analytic model was $g = 0.19$ (95% CI [0.11, 0.27], $p < .001$). 30.67% of the total variation in results could be attributed to heterogeneity between studies ($I^2_{\text{Level 3}}$) and 26.72% of the variation can be attributed to heterogeneity of multiple outcomes within studies ($I^2_{\text{Level 2}}$).

The results of the pooled SMD and moderator analyses are found in Table 3. Outcome Category was a significant moderator of SMD [$Q_m(3) = 17.56$, $p < .001$]. Table 4 summarizes pairwise post hoc tests with correction for multiple comparisons using the Benjamini-Hochberg method (i.e., false discovery rate). These tests revealed that intervention-related improvements in behavior change and quality of life were significantly greater than improvements in health indicators (SMD differences 0.34 and 0.32, respectively; ps = .01). Improvements in behavior change were also significantly greater than those for symptom change (SMD difference 0.12; $p = .02$). Outcomes did not significantly vary between intervention target [$Q_m(3) = 2.79$, $p = .42$]. Intervention type (treatment vs. prevention) was not a significant moderator of SMD [$Q_m(1) = 1.07$, $p = .30$] nor was integration model (co-located vs. integrated) [$Q_m(1) = 0.01$, $p = .92$].

Sensitivity analyses

When testing different levels of correlation between SMDs within studies ($\rho = 0.3$–0.7), there were no substantive differences in the results of the primary analyses (pooled SMD ranged from 0.18 to 0.19). We also tested for the presence of potentially influential and outlying SMD values by calculating Cook's distances for each outcome. The two studies (Kolko et al., 2012; Weersing et al., 2017) that used the Clinical Global Impression Scale (CGI) were identified by their Cook's distances for being influential. Upon removing these outcomes, the pooled SMD was negligibly smaller (SMD difference –0.04) and the moderator results were not substantially different.

Study design (randomized vs. non-randomized) approached significance for differences in SMD [$Q_m(1) = 2.86$, $p = .09$]. The randomized studies had a significant positive pooled SMD (SMD = 0.21, 95% CI [0.13, 0.29]) while the non-randomized studies had an SMD not statistically different from zero (SMD = 0.04, 95% CI [–0.32, 0.41]), but the two SMDs did not statistically differ from one another, likely due, in part, to the wide range of variability within the four non-randomized studies.

Racial and ethnic differences

Only two studies disaggregated outcome data to assess differences in clinical effectiveness across racial and ethnic groups

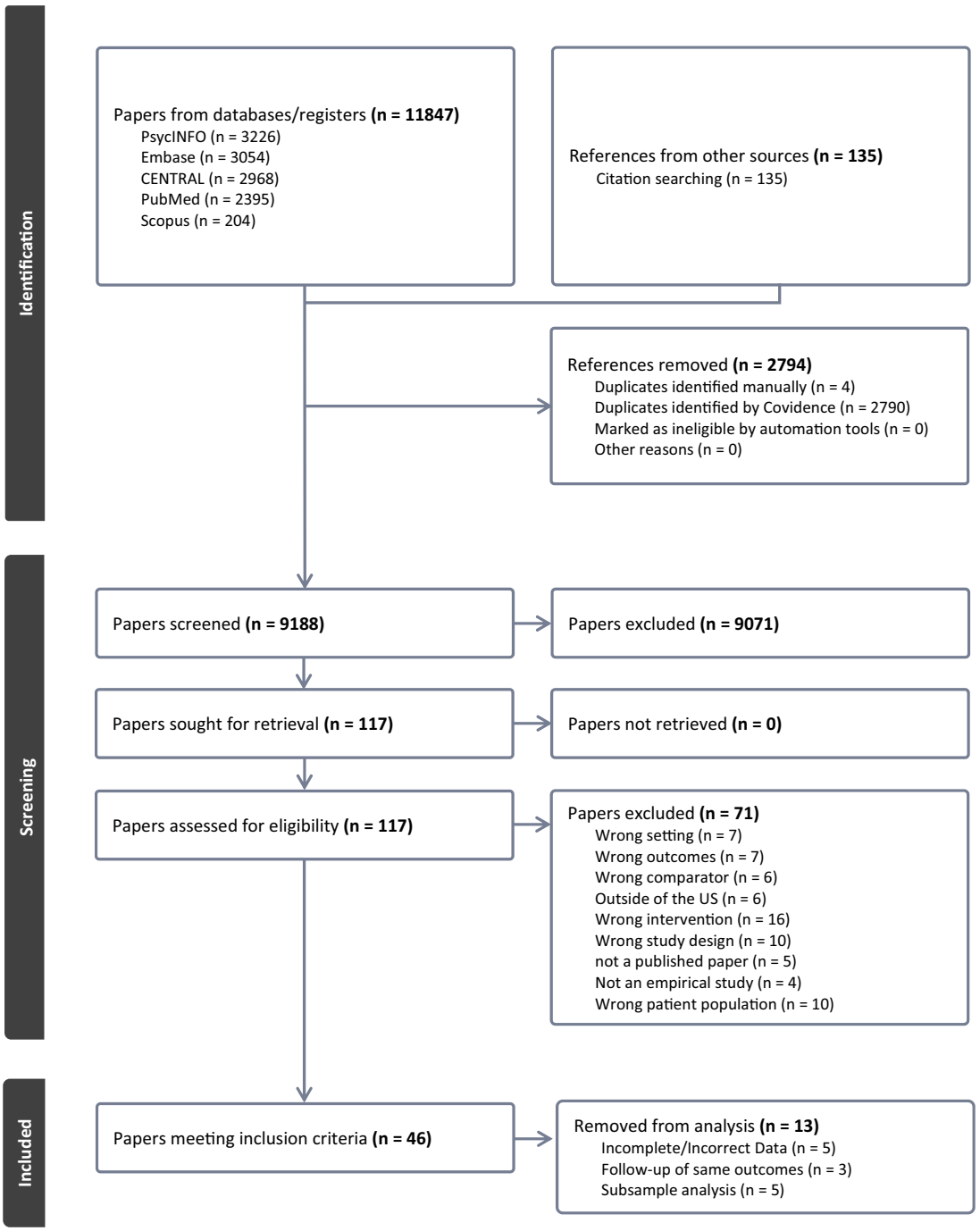

Figure 1. PRISMA flowchart.

which precluded inclusion of race/ethnicity as a moderator in the meta-analysis. Weersing et al. (2018) found that Hispanic youth (76.5% responding to treatment) had a larger response to co-located care than non-Hispanic participants (52.5% responding to treatment). Ngo et al. (2009) found that Black youth had significantly higher symptom reduction (SMD = 4.26) than White (SMD = 0.072) and Hispanic (SMD = 1.83) children in a follow-up analysis of the Asarnow et al. (2005) study. Both studies were deemed to have "some concerns" in ROB assessment due to some concerns with outcome measurement (note, ROB was not assessed for Ngo et al. as it involved the same dataset as Asarnow et al.).

Reporting of harms

Only Richardson et al. (2014) reported on adverse events. They did not find significant differences in psychiatric hospitalization between intervention and comparator groups and a trend toward more emergency room visits for control patients (10%) than intervention patients (2%). This study was evaluated as having "some concerns" due to outcome measurement through ROB assessment.

Certainty of evidence

Risk of bias

Of the 30 randomized controlled trial papers included in the meta-analysis, 1 paper was assessed as low risk of bias,

Table 2. Characteristics of included studies.

Study	Design	Trial type	Level of integration	Comparator	Interventionist	Intervention(s)	Intervention target	Outcome type(s)
Asarnow et al. (2005)	RCT	Treatment	Integrated	EUC	Masters/PhD in mental health/ nursing	CBT	Mental health	Quality of life Symptom
Bai et al. (2018)	RCT	Prevention	Co-located	UC	Masters level, psychologist	MI and CBT	Substance use	Behavior
Bauer et al. (2015)	RCT	Treatment	Co-located	UC	Behavioral pediatrician, medical resident	Psychoeducation	Mental health	Quality of life Symptom
Berkovits et al. (2010)	RCT	Treatment	Co-located	EUC	Psychology graduate student	Group PCIT	Mental health	Symptom
Briggs et al. (2012)	Quasi	Prevention	Integrated	UC	Psychologist	Healthy Steps	Development	Symptom
Briggs et al. (2014)	Quasi	Prevention	Integrated	UC	Psychologist	Healthy Steps	Development	Symptom
Canfield et al. (2015), Cates et al. (2016), Mendelsohn et al. (2011), and Weisleder et al. (2016)	RCT	Prevention	Co-located	EUC	Bachelor's level clinician	Video Interaction Project	Development	Behavior Quality of life Symptom
Clarke et al. (2005)	RCT	Treatment	Integrated	UC (with SSRI)	Master's level psychologist	CBT	Mental health	Quality of life Symptom
DeBar et al. (2012)	RCT	Treatment	Co-located	UC	Nutritionists, health educators, psychologist	Multicomponent (diet, activity, social, provider training)	Physical health	Health Quality of life
Gillham et al. (2006)	RCT	Prevention	Co-located	UC	Psychologist, social worker	CBT	Mental health	Symptom
Johnston et al. (2006)	C-RCT	Prevention	Integrated	UC	Master's level	Healthy Steps	Mental health	Behavior Symptom
Kolko et al. (2010)	RCT	Treatment	Co-located	EUC	Nurse	PMT, CBT, school consultation	Mental health	Quality of life Symptom
Kolko et al. (2012)	RCT	Treatment	Integrated	EUC	Master's level social worker, nurse	CBT, psychiatric consultation	Mental health	Symptom
Kolko et al. (2014)	C-RCT	Treatment	Integrated	EUC	Master's level social workers	CBT	Mental health	Quality of life Symptom
Linville et al. (2015)	RCT	Prevention	Co-located	EUC	"BH Interns"	Psychoeducation with supportive therapy	Mental health	Symptom
Mendelsohn et al. (2005) and Mendelsohn et al. (2007)	RCT	Prevention	Co-located	UC	Bachelor's level	Video Interaction Project	Development	Symptom Behavior

(continued)

Table 2. (continued)

Study	Design	Trial type	Level of integration	Comparator	Interventionist	Intervention(s)	Intervention target	Outcome type(s)
Minkovitz et al. (2001), Minkovitz et al. (2003), and Minkovitz et al. (2007)	RCT	Prevention	Co-located	UC	Nurses, early childhood educators, social workers	Healthy Steps	Mental health	Behavior
Perrin et al. (2014)	RCT	Treatment	Co-located	Waitlist	Psychologists, social workers	Group PMT	Mental health	Behavior Symptom
Power et al. (2014)	Quasi	Treatment	Integrated	EUC	Psychology fellows, community health partner	PMT, family-school consultation	Mental health	Behavior Quality of life Symptom
Richardson et al. (2014)	RCT	Treatment	Integrated	EUC	Master's level BH clinician	CBT	Mental health	Quality of life Symptom
Roby et al. (2021)	RCT	Prevention	Co-located	UC	Bachelor's level	PMT	Development	Behavior
Schilling et al. (2017)	RCT	Treatment	Co-located	Waitlist	NA	PMT	Mental health	Symptom
Smith et al. (2021)	RCT	Treatment	Co-located	UC	"Trained interventionist"	PMT	Physical health	Behavior Health
Sterling et al. (2018)	C-RCT	Treatment	Integrated	UC	Psychologist	MI, CBT, crisis management	Substance use and mental health	Symptom
Walton et al. (2013)	RCT	Treatment	Co-located	EUC	Research therapist	MI	Substance use	Quality of life Behavior
Walton et al. (2014)	RCT	Prevention	Co-located	EUC	Research therapist	MI	Substance use	Behavior
Weersing et al. (2017)	RCT	Treatment	Co-located	EUC	Master's level therapist	BT	Mental health	Symptom

Note. BH = behavioral health; BT = behavioral therapy; CBT = cognitive behavioral therapy; EUC = enhanced usual care; MI = motivational interviewing; PMT = parent management training; RCT = randomized controlled trial; C-RCT = cluster randomized controlled trial; Quasi = quasi-experimental; UC = usual care.

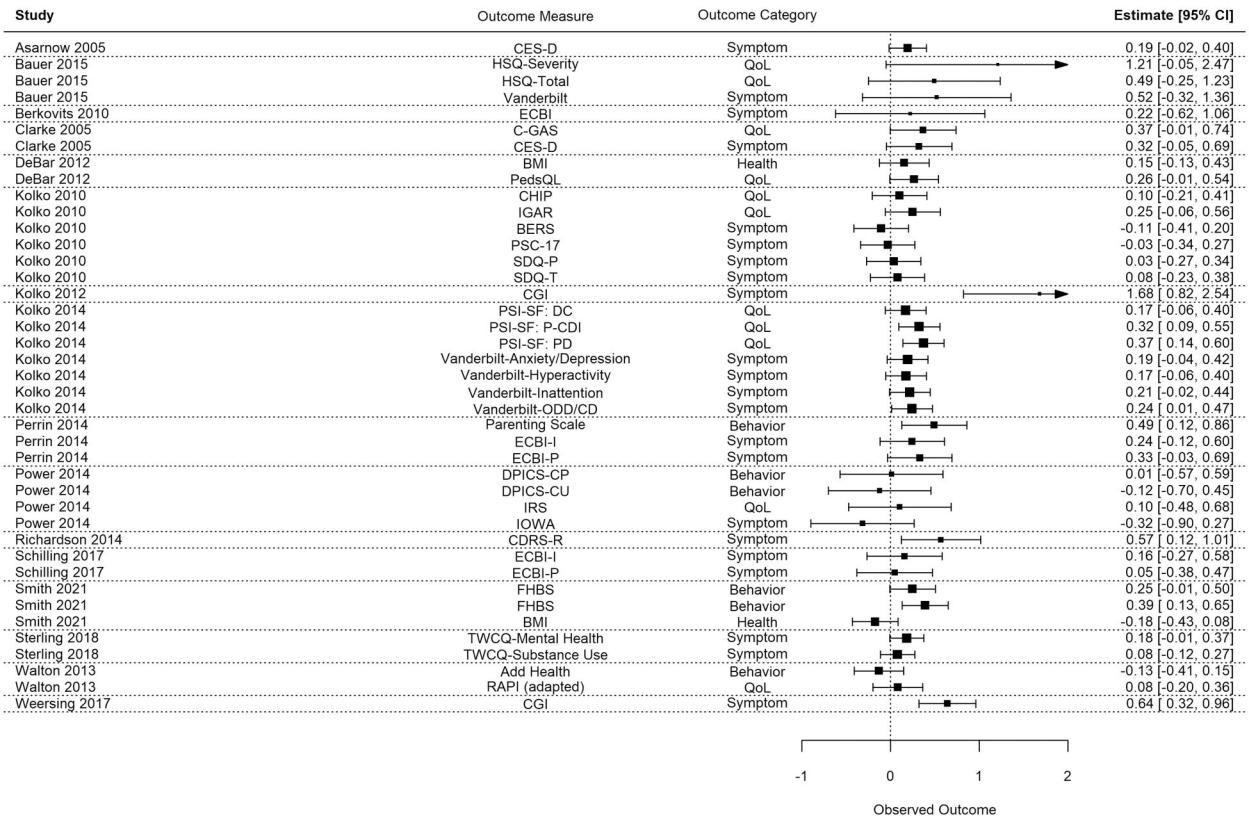

Figure 2. Treatment forest plot.

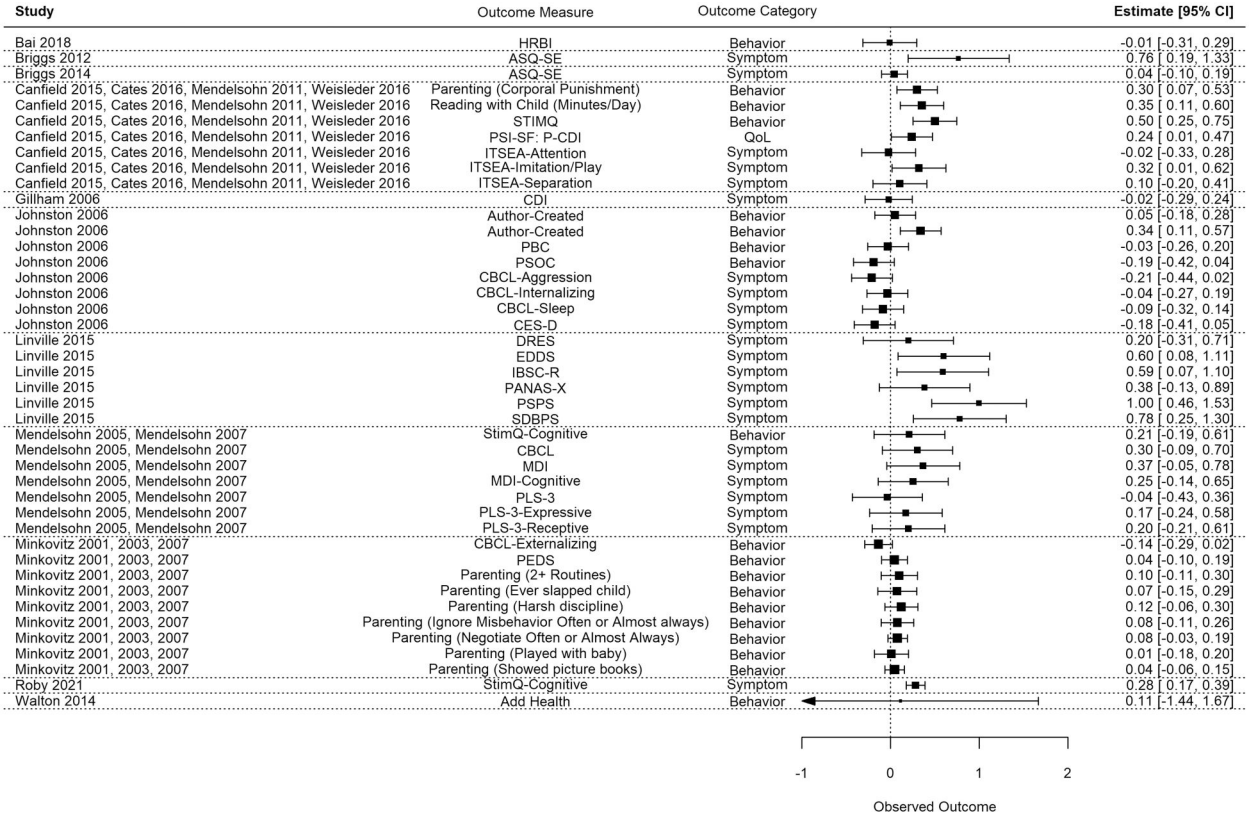

Figure 3. Prevention forest plot.

Table 3. Summary of meta-analysis results

Model	SMD	95% confidence interval		K	Q_m	df	p
		Lower limit	Upper limit				
Overall pooled SMD	0.19	0.11	0.27	27			<.001*
Moderator analyses							
Outcome type[a]					17.56	3	<.001*
Behavior change	0.29	0.15	0.43	10			<.001*
Health indicator	−0.05	−1.41	1.31	2			.83
Quality of life	0.27	0.18	0.36	8			<.001*
Symptom change	0.17	0.07	0.27	21			.003*
Intervention target[a]					2.79	3	.42
Development	0.23	0.07	0.39	6			.02*
Mental health	0.21	0.08	0.34	16			.005*
Physical health	0.18	−0.18	0.54	2			.10
Substance use	0.03	−0.16	0.23	4			.54
Intervention type					1.07	1	.30
Treatment	0.22	0.12	0.33	16			<.001*
Prevention	0.15	0.01	0.28	11			.04*
Integration model					0.01	1	.92
Co-located	0.19	0.09	0.29	17			.001*
Integrated	0.20	0.03	0.36	10			.03*
Sensitivity analyses							
Study design					2.86	1	.09
Randomized	0.21	0.13	0.29	23			<.001*
Non-randomized	0.04	−0.32	0.41	4			.66
Risk of bias					0.11	1	.74
High risk	0.21	−0.02	0.45	6			.07
Low/SOME Risk	0.18	0.09	0.28	21			<.001*

Note. Total sample size across studies: $N = 6{,}879$. SMD = standardized mean difference (Hedges' *g*). *K* = number of clusters (studies).
* $p < .05$.
[a] Total number of included studies for outcome type is greater than the grand total due to some studies including multiple outcome types.

Table 4. Outcome type post hoc comparisons

Outcome type comparison	SMD difference	SE	Z	p	$p_{corrected}$
Behavior change—Health indicator	0.34	0.11	3.10	.002	.01*
Behavior change—Quality of life	0.02	0.06	0.32	.75	.75
Behavior change—Symptom change	0.12	0.05	2.60	.009	.02*
Health indicator—Quality of life	−0.32	0.11	−2.82	.005	.01*
Health indicator—Symptom change	−0.22	0.11	−1.93	.05	.06
Quality of life—Symptom change	0.10	0.05	1.994	.05	.06

Note. Post hoc comparisons corrected using the Benjamini–Hochberg procedure. SMD = standardized mean difference (Hedges' *g*).
* $p < .05$.

20 were assessed as having some concerns, and 9 were assessed as high risk of bias. All three non-randomized studies were rated as a moderate overall risk of bias. Moderator analysis found no significant difference between studies of high versus low/some concern [$Q_m(1) = 0.02$, $p = .90$]. See Supplementary Table 1 for full outcomes of risk of bias assessment.

The most common risk of bias was due to the difficulty blinding participants to intervention status resulting in potential bias in measurement of outcomes, a common problem in psychotherapy research (Munder & Barth, 2018). This was the only type of measurement bias risk identified. Excluding items assessing blinding of participants and assessors would have resulted in 21 papers being assessed as low risk, 3 as some risk, and 6 as high risk of bias.

Four papers had some concern for risk of bias in the randomization procedure. One study assigned siblings to the group of the youngest child in the family (Bauer et al., 2015), one found significant between group differences on baseline measures (Smith et al., 2021), and two did not provide sufficient information to evaluate (Johnston et al., 2006; Sterling et al., 2018).

Two studies demonstrated some concern for deviating from intended interventions. Mendelsohn et al. (2011) reported that several participants randomized to the intervention group received care elsewhere for 6 months. Sterling et al. (2018) noted spillover effects and differences in fidelity among providers within intervention arms may have influenced study results within their limitations section.

Two papers reported high risk of bias related to missing data and one had some concern. The two studies with high bias had differential rates of missing data between treatment and comparator that were not random and associated with differences in baseline assessments (Cates et al., 2016; Linville et al., 2015). Mendelsohn et al. (2007) had some missing data but no clear indication that it was missing based on the true value of the missing item.

We found two papers to be at high risk of bias and one with some concern for bias related to the selection of reported results. The two papers with high risk of bias

reported different measures than what was listed in the trial protocol as outcomes (Mendelsohn et al., 2011; Weisleder et al., 2016). Bauer et al.'s (2015) paper was found to have some concern due to not specifying a pre-identified assessment plan.

The non-randomized studies were also unable to be blinded; however, the overall bias rating of moderate would not have changed if eliminating the expectation of blinding. The Power et al. (2014) paper also had a moderate risk due to differences in timing with recruitment of patients and the Briggs et al.'s papers (2012, 2014) did not provide enough information to rate whether there were deviations from intended interventions.

Reporting biases

Visual inspection of the funnel plot (Supplementary Figure 1) revealed possible asymmetry. Egger's test was significant ($\beta_0 = 0.14$, 95% CI [0.03, 0.24], $t = 2.68$, $p = .02$) suggesting a possible publication bias.

Discussion

In this multilevel meta-analysis, we examined the growing body of evidence for IPC interventions in pediatric primary care settings. An updated meta-analysis was warranted given recent clinical trials, methodological advances allowing for the testing of multiple outcomes nested within studies, and a continued emphasis on integrating medical and BH services in U.S. policy. To reflect the breadth and variety of IPC practice, we chose to include studies of interventions that ranged considerably in their targets (i.e., development, mental health, substance use, physical health), intensity (e.g., dose of treatment), and model of delivery (i.e., co-located and integrated models).

Replicating the main finding of Asarnow et al.'s (2015) previous meta-analysis, we found that IPC interventions were generally superior to usual care and enhanced usual care comparators when averaging across outcome types. Within outcome types, we found significant small-to-medium effects in studies targeting behavior change and quality of life. Smaller (but still significant) effects were found for symptom change. Only two included studies reported Health Indicators (Body Mass Index), for which we found no discernible effect. Unlike Asarnow et al. (2015), we did not find significant differences between prevention and treatment trials. We did not find differences in effectiveness across levels of integration which is similar to Asarnow et al. (2015) finding that collaborative care models were not significantly more effective than other models of integration in pediatrics. Relative to the overall effect size obtained by Asarnow et al. (2015; $d = 0.32$), we detected slightly smaller effect sizes. There are several possible reasons for this, including a slightly more stringent set of inclusion criteria that resulted in the exclusion of several studies with very large effect sizes in our study that were included in Asarnow et al. (2015) and our inclusion of studies of physical health interventions. We also used a multilevel analysis that allowed us to analyze multiple outcome domains per study as well as within-study (e.g., outcome type) and between-study (e.g., collaborative model) moderators. Our analysis revealed significant heterogeneity at both a within-study and between-study level. Furthermore, we were able to explain a portion of the within-study effects in terms of differences in outcome type. Multilevel meta-analytic methods represent an important development that allows for more nuanced and accurate assessment of outcomes than traditional approaches of selecting a single effect or averaging across multiple effect sizes (Cheung, 2019).

Limitations of the evidence base

The pediatric IPC trials literature is still relatively nascent, making synthesis challenging. This is particularly true given the considerable variation in the interventions studied. Indeed, IPC itself is best conceptualized as a vehicle for many potential interventions, rather than a discrete intervention in and of itself. A further complication of the relatively small, heterogeneous literature is the inconsistency with which authors define their overall integration models and the components of those models. This is partially a product of terminological evolution and imprecision, such that terms like "collaborative" and "integrated" are often used interchangeably and may function as both umbrella terms and labels for specific models of care. As such, it is difficult to reliably categorize interventions into particular models of integration, and synthesis is limited to the relatively broad evaluation of level of integration (i.e., "co-located" and "integrated"), which results in significant heterogeneity with regard to model components, interventions used, interventionists, outcomes types, and measurement.

While there was some evidence of differential effectiveness with Hispanic (Weersing et al., 2017) and Black (Ngo et al., 2009) youth demonstrating greatest benefit, few studies disaggregate outcomes by demographic characteristics making it difficult to assess equity of effectiveness. Further, although a number of IPC trials have been conducted in traditionally underserved communities with positive results (e.g., Asarnow et al., 2005; Mendelsohn et al., 2011; Power et al., 2014; Roby et al., 2021), studies often fail to report on the racial, ethnic, and socioeconomic composition of their samples in a manner that allows for robust synthesis. We could only extract race and ethnicity data from approximately 70% of studies and descriptions of samples were poor making it difficult to describe our meta-analytic sample with appropriate detail, accuracy, and inclusivity (APA, 2023). Even when samples are well described, it is typically unclear whether they are representative of their target population (Callejo-Black et al., 2020), so equity of access to IPC services remains an open question.

Our meta-analysis included only two studies (DeBar et al., 2012; Smith et al., 2021) that targeted physical health conditions. The goal of IPC is to improve the population health and well-being of all patients, not just those with psychosocial concerns (Reiter et al., 2018). Given significant rates of medical morbidity and associated impairment in youth, and the availability of evidence-based interventions for many pediatric populations (Palermo, 2014), the impact of IPC on physical health outcomes is understudied. Promising pilot studies have initially evaluated IPC interventions focused on obesity prevention (Rybak et al., 2023) and sleep problems (Williamson et al., 2022), but research focused on physical health outcomes remains relatively rare. Risk of bias ratings suggested a relatively low quality of studies; only one paper was rated low risk of bias. Heightened risk of bias was most commonly due to a dependence on self-reported or caregiver-reported outcomes measures (e.g., rating scales, interviews) and the inability to blind participants to intervention assignment. Although this often resulted in at least "some concern"

for bias, it should be noted that a previous meta-analysis on effectiveness of non-integrated psychotherapy found that outcomes reported by blinded outcome assessors typically resulted in higher effect sizes than self-reported outcomes suggesting that this bias may not lead to an overestimation of effect size (Cuijpers et al., 2010). Further, risk of bias results are impacted by the assessment tools used (Losilla et al., 2018), and calls have been made for greater emphasis on "selective outcome reporting" and "treatment implementation" over "blinding of personnel and participants" in psychotherapy outcome research (Munder & Barth, 2018). Thus, while the current analysis suggests "some concern" for bias in most included trials, the primary driver for this was due to bias risks that are inherent in psychological/behavioral health trials.

Limitations of the review process

Although this multilevel meta-analysis has numerous strengths, it is not without limitations. We limited our analyses to only those outcomes identified as primary by the author and there may be important secondary effects that were missed such as the impact of IPC on additional clinical outcomes, access, utilization, or cost. We observed significant heterogeneity in comparators (usual care, waitlist, and multiple different enhancements to usual care) and this can significantly impact effect sizes (Black et al., 2020; de Bruin et al., 2009, 2010). We also observed significant heterogeneity in the primary endpoint used, ranging from immediately following a single session intervention to 3 years after intervention completion. Further, errors or incomplete reporting of data prevented extraction from five studies (10% of the studies meeting inclusion criteria). Greater standardization of comparators, interventions and interventionists, definitions, and measurement of outcomes, as well as endpoint would help future attempts at synthesis. Lastly, many traditional ways of evaluating for publication bias (e.g., trim-and-fill method) have not yet been applied to multilevel or multivariate meta-analytic methods (Shi & Lin, 2019). More methodological work is needed to validate methods of evaluating for the risk of publication bias in more complex data structures.

Implications for practice, policy, and future research

Practice

The results support the practice of pediatric IPC, broadly, and highlight the potential varied roles of BH professionals in primary care across the span of child development. Participants in the reviewed studies ranged in age from birth to early adulthood, and interventions included primary, secondary, and tertiary levels of prevention and treatment. The results did not identify any differential effects based on level of integration, target outcomes (except for health indicators), or type of intervention (i.e., prevention versus treatment). Prior observational research has found differences between models of integration in areas other than effectiveness, including rates of access (Chakawa et al., 2020) and physician perception of benefits (Germán et al., 2017). It should be noted that the most prominent models of IPC (e.g., CoCM, PCBH) were initially developed for adult patient populations and represented a minority of the studies included in this review. Pediatric primary care has a unique set of demands and goals, and it may be that existing models are not optimized for pediatric care, so practitioners should weigh the pros and cons of different model components. Due to the nascent state of the pediatric literature identified in this review, poor definitions and reporting of models and components, and lack of differential effects based on the level of integration, it is difficult to determine optimizing factors beyond the conclusion that children and families experience greater symptom reduction, improved parenting and health behaviors, and have higher quality of life when evidence-based interventions are implemented *within* primary care practices both preventatively and as treatment. In practice, available IPC resources are often limited relative to population need, and clinicians must consider the unique characteristics of their patient population, clinical settings, and personal competencies in determining what approach to care is most appropriate (American Psychological Association, Presidential Task Force on Evidence-Based Practice, 2006).

Policy

The results of this meta-analysis are congruent with increased policy emphasis on integrating behavioral and medical services (Bagalman et al., 2022). We did not find evidence of one model of integration being markedly more common or one level of integration outperforming others. Therefore, policy and payment should focus on funding pediatric primary care and behavioral health integration broadly. We found evidence that both IPC prevention and treatment models improve symptoms, health behaviors, and quality of life across both mental health and developmental domains, with positive indicators of improved equity. Thus, we argue that the full continuum of IPC services, from prevention to treatment of mental health and developmental outcomes, should be supported in policy and payment. For example, we found positive effects for behavior change from multiple studies targeting early childhood parenting as a preventive mechanism (Canfield et al., 2015; Cates et al., 2016; Johnston et al., 2006; Mendelsohn et al., 2011; Minkovitz et al., 2001, 2003, 2007; Weisleder et al., 2016). Such interventions are well-established for promoting child development (Jeong et al., 2021); however, traditional fee-for-service payment models are unlikely to support prevention services, as they typically require patients to carry an existing mental health diagnosis. Alternative models including but not limited to global payment systems are needed to incentivize the full value of integrated care (Miller et al., 2017; Tynan, 2016). This is an especially critical area for policy change, as early childhood services to promote healthy development are among the best investments in health care (Bruner & Hayes, 2023). In recent years, several state and federal policies have begun to better acknowledge the range of valuable services IPC can provide. For example, the Centers for Medicare & Medicaid Services (2020) have created new Current Procedural Terminology codes to capture IPC care coordination services that are not accounted for in traditional mental health codes, and states have passed laws allowing for pre-deductible behavioral health well-visits to be reimbursed (H.B. 303, DE. 2022). Until more universal coverage of these and other non-traditional services are available in the United States at a rate that fully supports such services, the full expansion and effectiveness of IPC in pediatrics will likely be hampered.

Future research

Following the results of Asarnow et al.'s (2015) meta-analysis, Stancin (2016) argued for an increased focus on IPC RCTs, with an emphasis on comparing the relative strengths

and weakness of different models of integration. Eight years later, the need remains much the same, but also extends beyond traditional RCTs (e.g., hybrid trial designs Curran et al., 2012; Smith et al., 2021 as an example in this review) and should include pragmatic science methods to generate "practice-based evidence" (Riley & Freeman, 2019), including using the electronic health records and clinical registries to facilitate efficient data collection tools that allow for systematic collection of IPC data (Hostutler & Ramtekkar, 2021; Jetelina et al., 2018).

Experimental efficacy is important, but insufficient for translating research to practice and impacting population health (Glasgow et al., 2003). While the extant literature clearly indicates that IPC is generally superior to usual care, this finding is largely unsurprising, as most psychosocial interventions that have been studied in primary care are derived from therapies that are well-established in other settings (e.g., CBT, PMT). To develop a more pragmatic science, the field must move beyond the general question of, "Does IPC work?" to ask, "What interventions, when delivered in which models of IPC best improve what outcomes, for whom, and under what conditions?"

To answer those questions, it is essential that IPC research reporting standards improve with respect to both the components of IPC interventions and the settings and populations they are intended for. As Yonek et al.'s (2020) attempted component analysis illustrates, IPC research has not been conducted or disseminated in a manner that allows for identification of its "active ingredients" beyond a few core features (e.g., co-location of a BH professional or care manager). To determine the relative contributions of different components of IPC to care outcomes, they must be adequately defined and measured. While there have been attempts to operationally define unique models of integration (Funderburk et al., 2021; Reiter et al., 2018; Unützer et al., 2013), establish a unified lexicon (Peek, 2013), and develop objective measures of integration (Rose et al., 2023), the pediatric literature remains disjointed in its treatment of IPC as an independent variable. Further work is needed to better understand and taxonomize the existing pediatric literature and set forth clear standards of reporting IPC research. At a minimum, research on clinical interventions should always specify the model in which the intervention is delivered by including the name (when applicable) and describing the key components.

As noted above, the demographic characteristics of participants (and non-participants) in IPC trials are often poorly described, both in terms of study enrollment and intervention outcomes. While there has been some research into family preferences (Hails et al., 2023; Mehus et al., 2019; Riley et al., 2019, 2022; Williamson et al., 2020; Zimmermann et al., 2021), more work is needed to incorporate the voices of patients and families in IPC outcomes research. In addition to better understanding the patient populations that are most likely to benefit from IPC, more attention to the characteristics in which IPC is intended to be implemented is needed. To have wide impact, IPC interventions must be acceptable, feasible, and sustainable for integrated primary care teams, patients, and healthcare systems in non-research contexts, but few trials report on factors related to intervention adoption or implementation (Callejo-Black et al., 2020). The Kolko et al.'s (2010, 2012, 2014) studies within this review serve as positive examples for describing the primary care context and implementation indicators such as participation rate. In fact, a registered future trial (NCT04946253) suggests that Kolko et al. are seeking to better test and understand different implementation strategies for integrated primary care.

Conclusion

While much remains to be learned, the practice of pediatric IPC appears established as an effective practice. Future work should focus on the efficacy of specific, well-defined models of IPC and how they affect specified outcomes in discrete, well-defined populations. As the U.S. healthcare landscape continues to evolve, IPC researchers should strive to strike a balance between pragmatic work that accounts for current "real-world" conditions in most practice settings and producing innovative data to inform policy and change those conditions for the better.

Supplementary material

Supplementary material is available online at *Journal of Pediatric Psychology* (https://academic.oup.com/jpepsy/).

Author contributions

Cody Allen Hostutler (Conceptualization [lead], Data curation [lead], Formal analysis [supporting], Investigation [lead], Methodology [equal], Project administration [lead], Resources [equal], Software [equal], Supervision [equal], Validation [equal], Visualization [supporting], Writing—original draft [equal], Writing—review & editing [equal]), Jeffrey Shahidullah (Conceptualization [equal], Data curation [equal], Formal analysis [supporting], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [supporting], Resources [equal], Writing—original draft [equal], Writing—review & editing [equal]), Teryn Bruni (Conceptualization[equal], Data curation [equal], Investigation [equal], Methodology [equal], Validation [equal], Writing—original draft [supporting], Writing—review & editing [equal]), Kevin George Stephenson (Conceptualization[equal], Data curation [equal], Formal analysis [lead], Investigation [equal], Methodology [lead], Resources [equal], Software [equal], Validation [equal], Visualization [lead], Writing—original draft [equal], Writing—review & editing [equal]), Leah Utset (Conceptualization[equal], Data curation [equal], Formal analysis [supporting], Investigation [supporting], Methodology [supporting], Writing—review & editing [equal]), Leah LaLonde (Conceptualization[supporting], Data curation [equal], Investigation [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Jennifer A. Mautone (Conceptualization [equal], Data curation [equal], Formal analysis [supporting], Investigation [equal], Methodology [equal], Project administration [supporting], Resources [equal], Writing—original draft [equal], Writing—review & editing [equal]), Tiffany Rybak (Conceptualization[equal], Data curation [equal], Formal analysis [supporting], Investigation [equal], Methodology [equal], Project administration [supporting], Resources [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Chimereodo Okoroji (Conceptualization[supporting], Data curation [equal], Investigation [equal], Methodology [equal],

Writing—original draft [supporting], Writing—review & editing [equal]), Kurt Freeman (Conceptualization[equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [supporting], Writing—original draft [equal], Writing—review & editing [equal]), and Andrew Riley (Conceptualization[equal], Data curation [equal], Formal analysis [supporting], Investigation [equal], Methodology [equal], Project administration [lead], Supervision [lead], Validation [supporting], Writing—original draft [equal], Writing—review & editing [equal])

Acknowledgments

We would like to thank Katrice Stallworth for her support with final organization, completing tables, and references.

Funding

No funding was received to conduct this study.

Conflicts of interest: Authors have no conflicts of interests or disclosures.

References

American Psychological Association. (2023). *Inclusive language guide* (2nd ed.). https://www.apa.org/about/apa/ equity-diversity-inclusion/language-guidelines.pdf

American Psychological Association, Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285.

Asarnow, J. R., Jaycox, L. H., Duan, N., LaBorde, A. P., Rea, M. M., Murray, P., Anderson, M., Landon, C., Tang, L., & Wells, K. B. (2005). Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: A randomized controlled trial. *JAMA*, *293*(3), 311–319. https://doi.org/10.1001/jama.293.3.311

Asarnow, J. R., Jaycox, L. H., Tang, L., Duan, N., LaBorde, A. P., Zeledon, L. R., & Wells, K. B. (2009). Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. *American Journal of Psychiatry*, *166*(9), 1002-1010.

Asarnow, J. R., Rozenman, M., Wiblin, J., & Zeltzer, L. (2015). Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: A meta-analysis. *JAMA Pediatrics*, *169*(10), 929–937.

Bagalman, E., Dey, J., Jacobus-Kantor, L., Stoller, B., West, K., Radel, L., Schreier, A., Rousseau, M., Blanco, M., Creedon, T. B., Nye, E., Ali, M., Dubenitz, J. M., Schwartz, D., White, J. O., Swenson-O'Brien, A. J., Oberlander, S., Burnszynski, J., Lynch-Smith, M., … Haffajee, R. L. (2022). *HHS roadmap for behavioral health integration (Issue Brief)*. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. https://www.aspe.hhs.gov/reports/hhs-roadmap-behavioral-health-integration

Bai, S., Zeledon, L. R., D'Amico, E. J., Shoptaw, S., Avina, C., LaBorde, A. P., Anderson, M., Fitzpatrick, O. M., & Asarnow, J. R. (2018). Reducing health risk behaviors and improving depression in adolescents: A randomized controlled trial in primary care clinics. *Journal of Pediatric Psychology*, *43*(9), 1004–1016. https://doi.org/10.1093/jpepsy/jsy048

Bailey, Z. D., Krieger, N., Agénor, M., Graves, J., Linos, N., & Bassett, M. T. (2017). Structural racism and health inequities in the USA: Evidence and interventions. *Lancet (London, England)*, *389* (10077), 1453–1463.

Bauer, N. S., Szczepaniak, D., Sullivan, P. D., Mooneyham, G., Pottenger, A., Johnson, C. S., & Downs, S. M. (2015). Group visits to improve pediatric attention-deficit hyperactivity disorderc chronic care management. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *36*(8), 553–561. https://doi.org/10.1097/DBP.0000000000000207

Berkovits, M. D., O'Brien, K. A., Carter, C. G., & Eyberg, S. M. (2010). Early identification and intervention for behavior problems in primary care: A comparison of two abbreviated versions of parent-child interaction therapy. *Behavior Therapy*, *41*(3), 375–387. https://doi.org/10.1016/j.beth.2009.11.002

Black, N., Eisma, M. C., Viechtbauer, W., Johnston, M., West, R., Hartmann-Boyce, J., Michie, S., & de Bruin, M. (2020). Variability and effectiveness of comparator group interventions in smoking cessation trials: A systematic review and meta-analysis. *Addiction (Abingdon, England)*, *115*(9), 1607–1617. https://doi.org/10.1111/add.14969

Brent, D. A., Porta, G., Rozenman, M. S., Gonzalez, A., Schwartz, K. T., Lynch, F. L., & Weersing, V. R. (2020). Brief behavioral therapy for pediatric anxiety and depression in primary care: A follow-up. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(7), 856-867.

Briggs, R. D., Silver, E. J., Krug, L. M., Mason, Z. S., Schrag, R. D., Chinitz, S., & Racine, A. D. (2014). Healthy Steps as a moderator: The impact of maternal trauma on child social-emotional development. *Clinical Practice in Pediatric Psychology*, *2*(2), 166–175.

Briggs, R. D., Stettler, E. M., Silver, E. J., Schrag, R. D., Nayak, M., Chinitz, S., & Racine, A. D. (2012). Social-emotional screening for infants and toddlers in primary care. *Pediatrics*, *129*(2), e377-384–e384. https://doi.org/10.1542/peds.2010-2211

Bruner, C., & Hayes, M. (2023). *Federal health policy requires child health investment focus*. Health Affairs Forefront. https://www.healthaffairs.org/do/10.1377/forefront.20230920.430993

Burkhart, K., Asogwa, K., Muzaffar, N., & Gabriel, M. (2020). Pediatric integrated care models: a systematic review. *Clinical Pediatrics*, *59*(2), 148–153. https://doi.org/10.1177/0009922819890004

Callejo-Black, A., Wagner, D. V., Ramanujam, K., Manabat, A. J., Mastel, S., & Riley, A. R. (2020). A systematic review of external validity in pediatric integrated primary care trials. *Journal of Pediatric Psychology*, *45*(9), 1039–1052. https://doi.org/10.1093/jpepsy/jsaa068

Canfield, C. F., Weisleder, A., Cates, C. B., Huberman, H. S., Dreyer, B. P., Legano, L. A., Johnson, S. B., Seery, A., & Mendelsohn, A. L. (2015). Primary care parenting intervention effects on use of physical punishment among low-income parents of toddlers. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *36*(8), 586–593.

Cates, C. B., Weisleder, A., Dreyer, B. P., Johnson, S. B., Vlahovicova, K., Ledesma, J., & Mendelsohn, A. L. (2016). Leveraging healthcare to promote responsive parenting: Impacts of the video interaction project on parenting stress. *Journal of Child and Family Studies*, *25* (3), 827–835. https://doi.org/10.1007/s10826-015-0267-7

Cates, C. B., Weisleder, A., Johnson, S. B., Seery, A. M., Canfield, C. F., Huberman, H., Dreyer, B. P., & Mendelsohn, A. L. (2018). Enhancing parent talk, reading, and play in primary care: Sustained impacts of the Video Interaction Project. *The Journal of pediatrics*, *199*, 49-56.

Caughy, M. O. B., Miller, T. L., Genevro, J. L., Huang, K. Y., & Nautiyal, C. (2003). The effects of Healthy Steps on discipline strategies of parents of young children. *Journal of Applied Developmental Psychology*, *24*(5), 517-534.

Caughy, M. O. B., Huang, K. Y., Miller, T., & Genevro, J. L. (2004). The effects of the Healthy Steps for Young Children Program: Results from observations of parenting and child development. *Early Childhood Research Quarterly*, *19*(4), 611-630.

Centers for Medicare & Medicaid Services. (2020, December). *Revisions to payment policies under the Medicare Physician Fee Schedule*, Quality Payment Program and Other Revisions to Part B for CY 2021 (CMS-1734-F). https://www.cms.gov/medicaremedicare-fee-service-paymentphysicianfeeschedpfs-federal-regulation-notices/cms-1734-f

Chakawa, A., Belzer, L. T., Perez-Crawford, T., & Brei, N. (2020). Which model fits? Evaluating models of integrated behavioral health care in addressing unmet behavioral health needs among underserved sociodemographic groups. *Evidence-Based Practice in Child and Adolescent Mental Health*, *5*(3), 251–270. https://doi.org/10.1080/23794925.2020.1796549

Cheung, M. W. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review*, *29*(4), 387–396. https://doi.org/10.1007/s11065-019-09415-6

Chisolm, D. J., Klima, J., Gardner, W., & Kelleher, K. J. (2009). Adolescent behavioral risk screening and use of health services. *Administration and Policy in Mental Health*, *36*(6), 374–380. https://doi.org/10.1007/s10488-009-0245-8

Clarke, G., Debar, L., Lynch, F., Powell, J., Gale, J., O'Connor, E., Ludman, E., Bush, T., Lin, E. H., Von Korff, M., & Hertert, S. (2005). A randomized effectiveness trial of brief cognitive-behavioral therapy for depressed adolescents receiving antidepressant medication. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(9), 888–898. https://www.ncbi.nlm.nih.gov/pubmed/16113617

Copeland, W. E., Shanahan, L., Hinesley, J., Chan, R. F., Aberg, K. A., Fairbank, J. A., van den Oord, E. J. C. G., & Costello, E. J. (2018). Association of childhood trauma exposure with adult psychiatric disorders and functional outcomes. *JAMA Network Open*, *1*(7), e184493-e184493.

Covidence Software. (2019). *Veritas health innovation*. Melbourne, Australia. https://www.covidence.org/home

Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review*, *30*(6), 768–778.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care*, *50*(3), 217–226. https://doi.org/10.1097/MLR.0b013e3182408812

de Bruin, M., Viechtbauer, W., Hospers, H. J., Schaalma, H. P., & Kok, G. (2009). Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: Implications for the interpretation and comparison of intervention effects. *Health Psychology*, *28*(6), 668–674. https://doi.org/10.1037/a0015989

de Bruin, M., Viechtbauer, W., Schaalma, H. P., Kok, G., Abraham, C., & Hospers, H. J. (2010). Standard care impact on effects of highly active antiretroviral therapy adherence interventions: A meta-analysis of randomized controlled trials. *Archives of Internal Medicine*, *170*(3), 240–250.

DeBar, L. L., Stevens, V. J., Perrin, N., Wu, P., Pearson, J., Yarborough, B. J., Dickerson, J., & Lynch, F. (2012). A primary care-based, multicomponent lifestyle intervention for overweight adolescent females. *Pediatrics*, *129*(3), e611-620–e620. https://doi.org/10.1542/peds.2011-0863

Doleman, B., Freeman, S. C., Lund, J. N., Williams, J. P., & Sutton, A. J. (2020). Funnel plots may show asymmetry in the absence of publication bias with continuous outcomes dependent on baseline risk: Presentation of a new publication bias test. *Research Synthesis Methods*, *11*(4), 522–534.

Duong, M. T., Bruns, E. J., Lee, K., Cox, S., Coifman, J., Mayworm, A., & Lyon, A. R. (2021). Rates of mental health service utilization by children and adolescents in schools and other common service settings: A systematic review and meta-analysis. *Administration and Policy in Mental Health*, *48*(3), 420–439.

Funderburk, J. S., Polaha, J., & Beehler, G. P. (2021). What is the recipe for PCBH? Proposed resources, processes, and expected outcomes. *Families, Systems & Health*, *39*(4), 551–562.

Germán, M., Rinke, M. L., Gurney, B. A., Gross, R. S., Bloomfield, D. E., Haliczer, L. A., Colman, S., Racine, A. D., & Briggs, R. D. (2017). Comparing two models of integrated behavioral health programs in pediatric primary care. *Child and Adolescent Psychiatric Clinics of North America*, *26*(4), 815–828.

Gillham, J. E., Hamilton, J., Freres, D. R., Patton, K., & Gallop, R. (2006). Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology*, *34*(2), 203–219. https://doi.org/10.1007/s10802-005-9014-7

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, *93*(8), 1261–1267.

H.B. 303 (DE. 2022). 151st General Assembly 2021–2022.

Hacker, K. A., Myagmarjav, E., Harris, V., Suglia, S. F., Weidner, D., & Link, D. (2006). Mental health screening in pediatric practice: Factors related to positive screens and the contribution of parental/personal concern. *Pediatrics*, *118*(5), 1896–1906. https://doi.org/10.1542/peds.2006-0026

Hadland, S. E., Yehia, B. R., & Makadon, H. J. (2016). Caring for lesbian, gay, bisexual, transgender, and questioning youth in inclusive and affirmative environments. *Pediatric Clinics of North America*, *63*(6), 955–969. https://doi.org/10.1016/j.pcl.2016.07.001

Hails, K. A., Wellen, B. C., Simoni, M., Gaultney, W. M., Petts, R. A., Hostutler, C. A., & Riley, A. R. (2023). Parents' preferences for primary care-based behavioral services and the COVID-19 pandemic: A mixed method study. *Journal of Pediatric Psychology*, *48*(11), 879–892.

Hardin, A. P., Hackell, J. M., Simon, G. R., Boudreau, A. D. A., Baker, C. N., & Barden, G. A.; Committee on Practice and Ambulatory Medicine. (2017). Age limit of pediatrics. *Pediatrics*, *140*(3).

Heath, B., Wise Romero, P., & Reynolds, K. (2013). *A review and proposed standard framework for levels of integrated healthcare*. SAMHSA-HRSA Center for Integrated Health Solutions.

Hodgkinson, S., Godoy, L., Beers, L. S., & Lewin, A. (2017). Improving mental health access for low-income children and families in the primary care setting. *Pediatrics*, *139*(1), e20151175. https://doi.org/10.1542/peds.2015-1175

Hostutler, C. A., & Ramtekkar, U. P. (2021). Development and functionality of a pediatric behavioral health registry for integrated and collaborative care models. *Families, Systems & Health*, *39*(2), 366–373. https://doi.org/10.1037/fsh0000618

Hostutler, C., Wolf, N., Snider, T., Butz, C., Kemper, A. R., & Butter, E. (2023). Increasing access to and utilization of behavioral health care through integrated primary care. *Pediatrics*, *152*(6) https://doi.org/10.1542/peds.2023-062514

Jeong, J., Franchett, E. E., Ramos de Oliveira, C. V., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS Medicine*, *18*(5), e1003602. https://doi.org/10.1371/journal.pmed.1003602

Jetelina, K. K., Woodson, T. T., Gunn, R., Muller, B., Clark, K. D., DeVoe, J. E., Balasubramanian, B. A., & Cohen, D. J. (2018). Evaluation of an electronic health record (EHR) tool for integrated behavioral health in primary care. *Journal of the American Board of Family Medicine*, *31*(5), 712–723. https://doi.org/https://doi.org/10.3122/jabfm.2018.05.180041

Johnston, B. D., Huebner, C. E., Anderson, M. L., Tyll, L. T., & Thompson, R. S. (2006). Healthy steps in an integrated delivery system: Child and parent outcomes at 30 months. *Archives of Pediatrics & Adolescent Medicine*, *160*(8), 793–800. https://doi.org/10.1001/archpedi.160.8.793

Kolko, D. J., Campo, J., Kilbourne, A. M., Hart, J., Sakolsky, D., & Wisniewski, S. (2014). Collaborative care outcomes for pediatric behavioral health problems: A cluster randomized trial. *Pediatrics*, *133*(4), e981–e992. https://doi.org/10.1542/peds.2013-2516

Kolko, D. J., Campo, J. V., Kelleher, K., & Cheng, Y. (2010). Improving access to care and clinical outcome for pediatric behavioral problems: A randomized trial of a nurse-administered intervention in primary care. *Journal of Developmental and Behavioral*

Pediatrics: JDBP, *31*(5), 393–404. https://doi.org/10.1097/DBP.0b013e3181dff307

Kolko, D. J., Campo, J. V., Kilbourne, A. M., & Kelleher, K. (2012). Doctor-office collaborative care for pediatric behavioral problems: A preliminary clinical trial. *Archives of Pediatrics & Adolescent Medicine*, *166*(3), 224–231. https://doi.org/https://doi.org/10.1001/archpediatrics.2011.201

Kolko, D. J., Hart, J. A., Campo, J., Sakolsky, D., Rounds, J., Wolraich, M. L., & Wisniewski, S. R. (2020). Effects of collaborative care for comorbid attention deficit hyperactivity disorder among children with behavior problems in pediatric primary care. *Clinical pediatrics*, *59*(8), 787-800.

Lavigne, J. V., LeBailly, S. A., Gouze, K. R., Cicchetti, C., Pochyly, J., Arend, R., & Binns, H. J. (2008). Treating oppositional defiant disorder in primary care: A comparison of three models. *Journal of Pediatric Psychology*, *33*(5), 449-461.

Lavigne, J. V., LeBailly, S. A., Gouze, K. R., Binns, H. J., Keller, J., & Pate, L. (2010). Predictors and correlates of completing behavioral parent training for the treatment of oppositional defiant disorder in pediatric primary care. *Behavior Therapy*, *41*(2), 198-211.

Linville, D., Cobb, E., Lenee-Bluhm, T., Lopez-Zeron, G., Gau, J. M., & Stice, E. (2015). Effectiveness of an eating disorder preventative intervention in primary care medical settings. *Behaviour Research and Therapy*, *75*, 32–39. https://doi.org/10.1016/j.brat.2015.10.004

Losilla, J. M., Oliveras, I., Marin-Garcia, J. A., & Vives, J. (2018). Three risk of bias tools lead to opposite conclusions in observational research synthesis. *Journal of Clinical Epidemiology*, *101*, 61–72.

Mason, M., Pate, P., Drapkin, M., & Sozinho, K. (2011). Motivational interviewing integrated with social network counseling for female adolescents: A randomized pilot study in urban primary care. *Journal of Substance Abuse Treatment*, *41*(2), 148-155.

Mason, E. B., Burkhart, K., & Lazebnik, R. (2019). Adolescent stress management in a primary care clinic. *Journal of Pediatric Health Care*, *33*(2), 178–185.

McLaughlin, K. A., Hatzenbuehler, M. L., & Keyes, K. M. (2010). Responses to discrimination and psychiatric disorders among Black, Hispanic, female, and lesbian, gay, and bisexual individuals. *American Journal of Public Health*, *100*(8), 1477–1484. https://doi.org/10.2105/AJPH.2009.181586

McLeigh, J. D., Malthaner, L., Winebrenner, C., & Stone, K. E. (2022). Paediatric integrated care in the primary care setting: A scoping review of populations served, models used and outcomes measured. *Child: Care, Health and Development*, *48*(5), 869–879. https://doi.org/10.1111/cch.13000

Mehus, C. J., Kazlauskaite, V., Colianni, S., & Borowsky, I. W. (2019). Parents' interest in resources to address their child's behavioral health through primary care. *Families, Systems & Health*, *37*(3), 244–248. https://doi.org/10.1037/fsh0000429

Mendelsohn, A. L., Dreyer, B. P., Flynn, V., Tomopoulos, S., Rovira, I., Tineo, W., Pebenito, C., Torres, C., Torres, H., & Nixon, A. F. (2005). Use of videotaped interactions during pediatric well-child care to promote child development: A randomized, controlled trial. *Journal of Developmental & Behavioral Pediatrics*, *26*(6), 461–441. https://www.ncbi.nlm.nih.gov/pubmed/15718881

Mendelsohn, A. L., Huberman, H. S., Berkule, S. B., Brockmeyer, C. A., Morrow, L. M., & Dreyer, B. P. (2011). Primary care strategies for promoting parent-child interactions and school readiness in at-risk families: The Bellevue Project for Early Language, Literacy, and Education Success. *Archives of Pediatrics & Adolescent Medicine*, *165*(1), 33–41. https://doi.org/https://doi.org/10.1001/archpediatrics.2010.254

Mendelsohn, A. L., Valdez, P. T., Flynn, V., Foley, G. M., Berkule, S. B., Tomopoulos, S., Fierman, A. H., Tineo, W., & Dreyer, B. P. (2007). Use of videotaped interactions during pediatric well-child care: Impact at 33 months on parenting and on child development. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *28*(3), 206–212.

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989.

Merikangas, K. R., He, J-p., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., Georgiades, K., Heaton, L., Swanson, S., & Olfson, M. (2011). Service utilization for lifetime mental disorders in US adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(1), 32–45.

Miller, B. F., Ross, K. M., Davis, M. M., Melek, S. P., Kathol, R., & Gordon, P. (2017). Payment reform in the patient-centered medical home: Enabling and sustaining integrated behavioral health care. *The American Psychologist*, *72*(1), 55–68. https://doi.org/10.1037/a0040448

Miller-Matero, L. R., Khan, S., Thiem, R., DeHondt, T., Dubaybo, H., & Moore, D. (2019). Integrated primary care: Patient perceptions and the role of mental health stigma. *Primary Health Care Research & Development*, *20*, e48.

Minkovitz, C., Strobino, D., Hughart, N., Scharfstein, D., Guyer, B., & Healthy Steps Evaluation, T. (2001). Early effects of the healthy steps for young children program. *Archives of Pediatrics & Adolescent Medicine*, *155*(4), 470–479. https://doi.org/https://doi.org/10.1001/archpedi.155.4.470

Minkovitz, C. S., Hughart, N., Strobino, D., Scharfstein, D., Grason, H., Hou, W., Miller, T., Bishai, D., Augustyn, M., McLearn, K. T., & Guyer, B. (2003). A practice-based intervention to enhance quality of care in the first 3 years of life: The Healthy Steps for Young Children Program. *JAMA*, *290*(23), 3081–3091. https://doi.org/10.1001/jama.290.23.3081

Minkovitz, C. S., Strobino, D., Mistry, K. B., Scharfstein, D. O., Grason, H., Hou, W., Ialongo, N., & Guyer, B. (2007). Healthy Steps for Young Children: Sustained results at 5.5 years. *Pediatrics*, *120*(3), e658–e668. https://doi.org/10.1542/peds.2006-1205

Mufson, L., Dorta, K. P., Wickramaratne, P., Nomura, Y., Olfson, M., & Weissman, M. M. (2004). A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. *Archives of General Psychiatry*, *61*(6), 577–584.

Munder, T., & Barth, J. (2018). Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research*, *28*(3), 347–355.

National Survey of Children's Health. (n.d.). *NSCH 2018-19: Number of current or lifelong health conditions, nationwide, age in 3 groups website*. childhealthdata.org. Accessed February 24, 2021.

Ngo, V. K., Asarnow, J. R., Lange, J., Jaycox, L. H., Rea, M. M., Landon, C., Tang, L., & Miranda, J. (2009). Outcomes for youths from racial-ethnic minority groups in a quality improvement intervention for depression treatment. *Psychiatric Services*, *60*(10), 1357–1364.

Olufs, E. L., Valleley, R. J., Hembree, K. C., & Evans, J. H. (2016). Brief educational "curbside consultation": Impact on attention-deficit/hyperactivity disorder referrals in an integrated healthcare setting. *Families, Systems & Health*, *34*(3), 221–228.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery (London, England)*, *88*, 105906.

Palermo, T. M. (2014). Evidence-based interventions in pediatric psychology: Progress over the decades. *Journal of Pediatric Psychology*, *39*(8), 753–762.

Peek, C. J. (2013). Integrated behavioral health and primary care: A common language. In M. R. Talen & A. B. Valeras (Eds.), *Integrated behavioral health in primary care: Evaluating the evidence, identifying the essentials*. (pp. 9–31). Springer.

Perrin, E. C., Sheldrick, R. C., McMenamy, J. M., Henson, B. S., & Carter, A. S. (2014). Improving parenting skills for families of young children in pediatric settings: A randomized clinical trial. *JAMA Pediatrics*, *168*(1), 16–24. https://doi.org/10.1001/jamapediatrics.2013.2919

Power, T. J., Mautone, J. A., Marshall, S. A., Jones, H. A., Cacia, J., Tresco, K., Cassano, M. C., Jawad, A. F., Guevara, J. P., & Blum, N. J. (2014). Feasibility and potential effectiveness of integrated services for children with ADHD in urban primary care practices. *Clinical Practice in Pediatric Psychology*, *2*(4), 412–426.

Pustejovsky, J. (2023). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections*. R package version 0.5.10. https://CRAN.R-project.org/package=clubSandwich.

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science*, *23*(3), 425–438.

Reiter, J. T., Dobmeyer, A. C., & Hunter, C. L. (2018). The primary care behavioral health (PCBH) model: An overview and operational definition. *Journal of Clinical Psychology in Medical Settings*, *25*(2), 109–126.

Richardson, L. P., Ludman, E., McCauley, E., Lindenbaum, J., Larison, C., Zhou, C., Clarke, G., Brent, D., & Katon, W. (2014). Collaborative care for adolescents with depression in primary care: A randomized clinical trial. *JAMA*, *312*(8), 809–816. https://doi.org/10.1001/jama.2014.9259

Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. *Pediatrics*, *126*(6), 1117–1123.

Richman, E. L., Lombardi, B. M., & Zerden, L. D. (2020). Mapping colocation: Using national provider identified data to assess primary care and behavioral health colocation. *Families, Systems & Health*, *38*(1), 16–23.

Riley, A. R., & Freeman, K. A. (2019). Impacting pediatric primary care: Opportunities and challenges for behavioral research in a shifting healthcare landscape. *Behavior Analysis (Washington, D.C.)*, *19*(1), 23–38. https://doi.org/10.1037/bar0000114

Riley, A. R., Walker, B. L., Wilson, A. C., Hall, T. A., Stormshak, E. A., & Cohen, D. J. (2019). Parents' Consumer Preferences for Early Childhood Behavioral Intervention in Primary Care. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *40*(9), 669–678. https://doi.org/10.1097/DBP.0000000000000736

Riley, A. R., Walker, B. L., Ramanujam, K., Gaultney, W. M., & Cohen, D. J. (2022). A mixed-method investigation of parent perspectives on early childhood behavioral services in primary care. *The Journal of Behavioral Health Services & Research*, *49*(2), 134–148. https://doi.org/10.1007/s11414-021-09772-2

Roby, E., Miller, E. B., Shaw, D. S., Morris, P., Gill, A., Bogen, D., Rosas, J., Canfield, C. F., Hails, K. A., Wippick, H., Honoroff, J., Cates, C. B., Weisleder, A., Chadwick, K., Raak, C. D., & Mendelsohn, A. L. (2021). Improving parent-child interactions in pediatric health care: A two-site randomized controlled trial. *Pediatrics*, *147*(3) https://doi.org/10.1542/peds.2020-1799

Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*, *26*(2), 141.

Rose, G. L., Weldon, T. L., McEntee, M. L., Hitt, J. R., Kessler, R., Littenberg, B., Macchi, C. R., Martin, M. P., Mullin, D. J., & van Eeghen, C. (2023). Practice integration profile revised: Improving item readability and completion. *Families, Systems & Health*, *41*(2), 201–206.

Rybak, T. M., Modi, A. C., Mara, C. A., Herbst, R. B., Lauer, B. A., Burkhardt, M. C., & Stark, L. J. (2023). A pilot randomized trial of an obesity prevention program for high-risk infants in primary care. *Journal of Pediatric Psychology*, *48*(2), 123–133.

Schilling, S., French, B., Berkowitz, S. J., Dougherty, S. L., Scribano, P. V., & Wood, J. N. (2017). Child-adult relationship enhancement in primary care (PriCARE): A randomized trial of a parent training for child behavior problems. *Academic Pediatrics*, *17*(1), 53–60. https://doi.org/10.1016/j.acap.2016.06.009

Schwartz, K. T., Kado-Walton, M., Dickerson, J. F., Rozenman, M., Brent, D. A., Porta, G., & Weersing, V. R. (2023). Brief behavioral therapy for anxiety and depression in pediatric primary care: Breadth of intervention impact. *Journal of the American Academy of Child & Adolescent Psychiatry*, *62*(2), 230-243.

Shaffer, A., Lindhiem, O., & Kolko, D. (2017). Treatment effects of a primary care intervention on parenting behaviors: Sometimes it's relative. *Prevention Science*, *18*, 305-311.

Shahidullah, J. D., Carlson, J. S., Haggerty, D., & Lancaster, B. M. (2018). Integrated care models for ADHD in children and adolescents: A systematic review. *Families, Systems & Health*, *36*(2), 233–247. https://doi.org/10.1037/fsh0000356

Shahidullah, J. D., Hostutler, C. A., Coker, T. R., Allmon Dixson, A., Okoroji, C., & Mautone, J. A. (2023). Child health equity and primary care. *The American Psychologist*, *78*(2), 93–106. https://doi:10.1037/amp0001064

Shi, L., & Lin, L. (2019). The trim-and-fill method for publication bias: Practical guidelines and recommendations based on a large database of meta-analyses. *Medicine*, *98*(23), e15987.

Smith, E. E., Rome, L. P., & Freedheim, D. K. (1967). The clinical psychologist in the pediatric office. *The Journal of Pediatrics*, *71*(1), 48–51. https://doi.org/10.1016/s0022-3476(67)80229-4

Smith, J. D., Berkel, C., Carroll, A. J., Fu, E., Grimm, K. J., Mauricio, A. M., Rudo-Stern, J., Winslow, E., Dishion, T. J., Jordan, N., Atkins, D. C., Narayanan, S. S., Gallo, C., Bruening, M. M., Wilson, C., Lokey, F., & Samaddar, K. (2021). Health behaviour outcomes of a family based intervention for paediatric obesity in primary care: A randomized type II hybrid effectiveness-implementation trial. *Pediatric Obesity*, *16*(9), e12780. https://doi.org/10.1111/ijpo.12780

Stancin, T., & Perrin, E. C. (2014). Psychologists and pediatricians: Opportunities for collaboration in primary care. *The American Psychologist*, *69*(4), 332–343. https://doi.org/10.1037/a0036046

Stancin T. (2016). Commentary: Integrated pediatric primary care: Moving from why to how. *Journal of pediatric psychology*, *41*(10), 1161–1164. https://doi.org/10.1093/jpepsy/jsw074

Sterling, S., Kline-Simon, A. H., Weisner, C., Jones, A., & Satre, D. D. (2018). Pediatrician and behavioral clinician-delivered screening, brief intervention and referral to treatment: Substance use and depression outcomes. *The Journal of Adolescent Health*, *62*(4), 390–396.

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., … Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical Research ed.)*, *355*, i4919.

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research ed.)*, *366*, l4898.

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634.

Trent, M., Dooley, D. G., Dougé, J., Cavanaugh, R. M., Jr, Lacroix, A. E., Fanburg, J., Rahmandar, M. H., Hornberger, L. L., Schneider, M. B., Yen, S., Chilton, L. A., Green, A. E., Dilley, K. J., Gutierrez, J. R., Duffee, J. H., Keane, V. A., Krugman, S. D., McKelvey, C. D., Linton, J. M., Nelson, J. L., Mattson, G., … Wallace, S. B.; Section on Adolescent Health (2019). The impact of racism on child and adolescent health. *Pediatrics*, *144*(2), Article e20191765. https://doi.org/10.1542/peds.2019-1765

Tynan, W. D. (2016). Commentary: Integrated primary care psychology evolving with the affordable care act. *Journal of Pediatric Psychology*, *41*(10), 1165–1167. https://doi.org/10.1093/jpepsy/jsw086

Unützer, J., Harbin, H., Schoenbaum, M., & Druss, B. (2013). The collaborative care model: An approach for integrating physical and mental health care in Medicaid health homes [Issue Brief]. Center for Health Care Strategies, *Health Home Information Resource Center*. https://www.chcs.org/media/HH_IRC_Collaborative_Care_Model__052113_2.pdf

Varni, J. W., Burwinkle, T. M., & Seid, M. (2005). The PedsQL™ as a pediatric patient-reported outcome: Reliability and validity of the PedsQL™ measurement model in 25,000 children. *Expert Review of Pharmacoeconomics & Outcomes Research*, *5*(6), 705–719.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Walton, M. A., Bohnert, K., Resko, S., Barry, K. L., Chermack, S. T., Zucker, R. A., Zimmerman, M. A., Booth, B. M., & Blow, F. C. (2013). Computer and therapist based brief interventions among cannabis-using adolescents presenting to primary care: One year outcomes. *Drug and Alcohol Dependence*, *132*(3), 646–653. https://doi.org/10.1016/j.drugalcdep.2013.04.020

Walton, M. A., Resko, S., Barry, K. L., Chermack, S. T., Zucker, R. A., Zimmerman, M. A., Booth, B. M., & Blow, F. C. (2014). A randomized controlled trial testing the efficacy of a brief cannabis universal prevention program among adolescents in primary care. *Addiction (Abingdon, England)*, *109*(5), 786–797. https://doi.org/10.1111/add.12469

Wang, S., Li, Q., Lu, J., Ran, H., Che, Y., Fang, D., Liang, X., Sun, H., Chen, L., Peng, J., Shi, Y., & Xiao, Y. (2023). Treatment rates for mental disorders among children and adolescents: A systematic review and meta-analysis. *JAMA Network Open*, *6*(10), e2338174. https://doi.org/10.1001/jamanetworkopen.2023.38174

Weersing, V. R., Brent, D. A., Rozenman, M. S., Gonzalez, A., Jeffreys, M., Dickerson, J. F., Lynch, F. L., Porta, G., & Iyengar, S. (2017). Brief behavioral therapy for pediatric anxiety and depression in primary care: A randomized clinical trial. *JAMA Psychiatry*, *74*(6), 571–578. https://doi.org/10.1001/jamapsychiatry.2017.0429

Weisleder, A., Cates, C. B., Dreyer, B. P., Berkule Johnson, S., Huberman, H. S., Seery, A. M., Canfield, C. F., & Mendelsohn, A. L. (2016). Promotion of Positive Parenting and Prevention of Socioemotional Disparities. *Pediatrics*, *137*(2), e20153239. https://doi.org/10.1542/peds.2015-3239

Whitney, D. G., & Peterson, M. D. (2019). US national and state-level prevalence of mental health disorders and disparities of mental health care use in children. *JAMA Pediatrics*, *173*(4), 389–391.

Williamson, A. A., Milaniak, I., Watson, B., Cicalese, O., Fiks, A. G., Power, T. J., Barg, F. K., Beidas, R. S., Mindell, J. A., & Rendle, K. A. (2020). Early childhood sleep intervention in urban primary care: Caregiver and clinician perspectives. *Journal of Pediatric Psychology*, *45*(8), 933–945. https://doi.org/10.1093/jpepsy/jsaa024

Williamson, A. A., Okoroji, C., Cicalese, O., Evans, B. C., Ayala, A., Harvey, B., Honore, R., Kratchman, A., Beidas, R. S., Fiks, A. G., Power, T. J., & Mindell, J. A. (2022). Sleep Well! An adapted behavioral sleep intervention implemented in urban primary care. *Journal of Clinical Sleep Medicine: JCSM*, *18*(4), 1153–1166.

Wilson, D. B. (2023). *Practical meta-analysis effect size calculator*. Version 2023.11.27. https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

Wissow, L. S., Gadomski, A., Roter, D., Larson, S., Brown, J., Zachary, C., Bartlett, E., Horn, I., Luo, X., & Wang, M.-C. (2008). Improving child and parent mental health in primary care: A cluster-randomized trial of communication skills training. *Pediatrics*, *121*(2), 266–275.

Yonek, J., Lee, C. M., Harrison, A., Mangurian, C., & Tolou-Shams, M. (2020). Key components of effective pediatric integrated mental health care models. *JAMA Pediatrics*, *174*(5), 487–498. https://doi.org/10.1001/jamapediatrics.2020.0023

Zimmermann, M., O'Donohue, W., Zepeda, M., & Woodley, A. (2021). Examining caretaker attitudes towards primary prevention of pediatric behavioral health problems in integrated care. *The Journal of Behavioral Health Services & Research*, *48*(1), 120–132. https://doi.org/10.1007/s11414-020-09720-6

Journal of Pediatric Psychology, 2024, 00, 1–18
https://doi.org/10.1093/jpepsy/jsae038
Systematic Review