NATIONAL QUALITY FORUM
Driving measurable health improvements together

Measure Worksheet

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.
To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 3332
Corresponding Measures:
Measure Title: Psychosocial Screening Using the Pediatric Symptom Checklist-Tool (PSC-Tool)
Measure Steward: Massachusetts General Hospital
sp.02. Brief Description of Measure: Percentage of children from 3.00 to 17.99 years of age seen for a pediatric well child visit who have a Pediatric Symptom Checklist (PSC) Tool administered as a component of that visit.
1b.01. Developer Rationale: Psychosocial problems in children are common and treatable with prevalence estimates of about 12% of all children and adolescents (Gardner, Lucas, Kolko, & Campo, 2007; Kelleher et al., 1997; Murphy et al., 2016). Studies have shown that children with these problems are often unrecognized by their pediatricians (~50% of cases) (Kelleher et al., 1997) and that only a fraction of them receive treatment (Hacker et al., 2014b; Kelleher et al., 1997). Children with psychosocial problems are more likely to have poorer health, academic, behavioral, and social outcomes in both the short and long term (Murphy et al., 2015). Children who receive psychosocial screening as a part of pediatric well child visits are more likely to receive outpatient mental health services (Hacker et al., 2014a; Hacker et al., 2014b; Savageau et al., 2016) than are children who are not screened. As the dates of the studies just cited attest, it is only within the last three years that strong evidence documenting the relationship between psychosocial screening and increased mental health treatment has become available.
A series of RCT studies by Kolko and his associates have shown that pediatric outpatients with a wide range of problems who are found to be at risk when screened with the PSC and go on to receive pediatric office based mental health interventions have significantly lower mental health symptom scores and better functioning at immediate and longer term follow up than do similar outpatients randomized to treatment as usual (Kolko et al., 2014; Kolko, Campo, Kelleher, & Cheng, 2010). For these reasons, we believe that an increase in mental health treatment is the most appropriate (and a measurable) benchmark for assessing the positive impact of routine psychosocial screening. The logic model for screening in pediatrics is that more children will receive help, fewer children will develop mental, emotional, and behavioral disorders (Guzmán et al., 2015; Kieling et al., 2011), and more children who received help will enjoy better life outcomes (Kellam et al., 2014).
Requiring screening for psychosocial problems as part of routine well child care in pediatrics is one of the most frequently recommended ways to improve recognition and intervention for such problems

(Hacker et al., 2014a) and an increasing number of states (Massachusetts (Savageau et al., 2016)), insurers (Medicaid/EPSDT (Mann, 2013)), standard setting organizations (American Academy of Pediatrics (Foy, Kelleher, Laraque, & Health, 2010; Weitzman & Wegner, 2015)), blue ribbon panels (President's New Freedom Commission on Mental Health (Hogan, 2003) (Institute of Medicine (O´Connell, Boat, & Warner, 2009)), and advocacy organizations such as the Kennedy Forum (Fortney et al., 2015) and Mental Health America (http://www.mentalhealthamerica.net/positions/early-identification) have now required, endorsed, or recommended the principle of including a psychosocial screen as a part of every well child visit for children aged 3-17.

The PSC is probably the most frequently recommended and widely used tool for routine psychosocial screening in pediatrics (Semansky, Koyanagi, & Vandivort-Warren, 2003), with the Strengths and Difficulties Questionnaire (Goodman, Meltzer, & Bailey, 1998) and Child Behavior Checklist (Achenbach, 2009) instruments that are similar in many ways and also frequently mentioned and validated in this context. Many of the endorsements noted above include these three and/or a few other instruments.

The reference list is included in the attached appendix.

sp.12. Numerator Statement: Number of patients with documentation that the PSC tool was administered as part of the well child visit.

sp.14. Denominator Statement: Number of patients aged 3.00 to 17.99 seen for a pediatric well-child visit.

sp.16. Denominator Exclusions: No exclusions.

Measure Type: Process
sp.28. Data Source: Claims
sp.07. Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: 2018-05-16
Most Recent Endorsement Date: 5/16/2018

IF this measure is included in a composite, NQF Composite#/title:
IF this measure is paired/grouped, NQF#/title:
sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

Preliminary Analysis: Maintenance of Endorsement Measure

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a ***structure, process or intermediate outcome*** measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☐ **Yes** ☒ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☐ **Yes** ☒ **No**
- **Evidence graded?** ☐ **Yes** ☒ **No**

Summary of prior review in 2017

- This is a process measure utilizing claims data at the health plan level to assess the percentage of children who have a PSC-Tool administered during a pediatric well child visit.
- During the previous review in 2017, the developer provided evidence of studies demonstrating the feasibility and acceptability of the PSC-Tool as a clinical and research measure with diverse populations and on a statewide scale.
- The logic model indicated that this process measure will increase the likelihood of identifying, exploring, and dealing with psychosocial issues during a well-child visit.
- Several members of the Standing Committee expressed concern regarding lack of evidence but ultimately agreed that there were a series of randomized controlled trials linking screening with the PSC-Tool to improved outcomes for children who are identified to be at risk.

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

The developer presents the following evidence to support this measure:

- The developer cited the American Academy of Pediatrics (AAP) Task Force on Mental Health's recommendation for routine psychosocial screening as a part of all well child visits (AAP, 2021).
- The developer highlighted the Bright Futures/AAP Recommendations for Preventive Pediatric Health Care (Periodicity Schedule), which advocates for the inclusion of "psychosocial/ behavioral assessment" at every well child visit from birth to age 21.
- One hundred and twenty studies (citations are listed in Appendix 1) published in the past four years that support the measure; either used the PSC as a measure or reviewed it along with “other measures” (i.e., screening tools).
- The provided evidence identifies a conceptual model from screening to interventions to outcomes/ intermediate outcome. Traditionally, NQF prefers endorses screening, intermediate outcome and follow up when clinically indicated, rather than screening only measures.

Questions for the Committee:

- The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review.

Guidance from the Evidence Algorithm

Process measure with empirical evidence submitted but not systematically reviewed (Box 3) → empirical evidence without systematic review/grading of evidence (Box 7) → evidence summarized include all studies (box 8) → high certainty that benefits outweigh undesirable effects (box 9) → Moderate

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided data from several studies demonstrating a wide variation in the rates of mental health screening with formal, standardized tools as well as the possibility of improvement and the potential benefits of doing so:
  - Summary data from the Children's Behavioral Health Initiative (CBHI) Behavioral Health Screening Cumulative Quarterly Report. Measurement Period dates of service from 1/1/2008 to 3/31/2017. Data Source: Statewide data for all children.
    - The range in rates was from 14.2% to 71.9% with a statewide average of 62.8% and a standard deviation of 12.4%. (95% CI: 58.8% to 66.8%).
  - Distribution of screening rates demonstrating the wide range of rates of screening across states over nearly a decade.
    - The range in rates was from 8.21% to 85.65% with a median of 64.53%.
  - Statewide data retrieved from chart reviews supplemented by administrative claims data showing variation in screening rates by region.
    - The range in rates was from 1.5% to 88.9% with a standard deviation of 35.1%. (95% CI: 35.5% to 87.2%).
  - Differences in screening rates from four Massachusetts General Hospital (MGH) outpatient clinics.
    - The range in rates was from 9.4% to 91.7% with an average of 76.6% and a standard deviation of 38.5%. (95% CI: 38.8% to 100.0%).

Disparities

- The developer presents data from the CBHI BHSCQR which illustrates different screening rates by age group with very young (<3) children [56.44%] and older (>17) youth [35.04%] less likely to be screened than 3–17-year-olds [71.22%].
- The developer presents data from a chart review study exploring screening by demographics in a subsample of ~6000 visits from 2007, 2010 and 2012.
- The developer reports no significant disparities by race, ethnicity, or language.
- The developer noted that disparities by socioeconomic status cannot be meaningfully assessed in this sample since by CBHI is a program only for children with Medicaid and SES is confounded with insurance type.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- This is a maintenance measure. The developers provided a large amount of new medical evidence to support the use of this measure. This submission continues to support the use of this measure.
- Moderate- Maintenance measure. Updated evidence stronger.
- Evidence to support this measure is moderate. Multiple studies are cited, and the position of the American Academy of Pediatrics are cited. No systematic reviews have been done in this area. Screening may lead to lead to prevention and early intervention, which is cost effective, according to the measure documents. Data presented provides evidence that screening can be done. The connection of this measure to improving referrals which improve outcomes for children and families is not as strong. Additionally, 18% of well child care is provided by family physicians, and I am not certain that the evidence includes other than pedestrian practices, except where mandated.
- One hundred and twenty studies (citations are listed in Appendix 1) published in the past four years that support the measure; either used the PSC as a measure or reviewed it along with “other measures” (i.e., screening tools).
- Since a process measure the focus is on increasing the percentage of Psychosocial screening during Well Child Visits to improve behavioral health. The evidence for this measure in well-documented (difficult to review entire list with some many references). Evidence suggests measure is increasingly being used and the NQF measure would provide valuable data to improve pediatric care.
- While there is no systematic review of evidence nor is there a link between screening and improved mental health outcomes, there is evidence from multiple RCTs on improvement in intermediate outcomes and access to treatment with screening; multiple organizations endorse screening; there is more evidence then presented when the measure was initially endorsed
- While there is a substantial link between screening for mental health disorders and providing interventions, the focus of this measure is a singular instrument rather than the process of screening. The authors themselves admit there are other validated tools for the same task. Moreover, the focus on process, rather than outcomes tempers my enthusiasm for this measure.

- Process measure examining rate of completion of a pediatric mental/behavioral health screening tool at the time of a well-child visit. The evidence directly suggests that, when this occurs, the patient is more likely to be referred for evaluation to the appropriate MH/BH professional, appropriately diagnosed with a BH/MH disorder, AND to receive treatment. Tangentially, the evidence implies that when these outcomes occur that better BH/MH outcomes result. I know of no studies which need to be additionally considered at this time.

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- The developers provided longitudinal data since the initial submission to demonstrate persistent gaps in performance. There was evidence of regional variation and some evidence of disparities in measurement by subgroup.
- Moderate room for improvement especially with younger and older children. I was surprised to see no significant gaps in screening rates by race.
- There are systems that have been strong at screening, many of which are large hospital or academic centers. Since this measure was first reviewed in NQF - "there has been a geometric increase in screening". Data from those screening supports more referrals to mental health resources. Screening rates can be as low as 14% to as high as 70+%.
- The developer provided data from several studies demonstrating a wide variation in the rates of mental health screening with formal, standardized tools as well as the possibility of improvement and the potential benefits of doing so. Rate ranged from 14.2% to 71.9% indicating a wide gap in performance.
- Currently appears to be used for Medicaid, although some references to other populations. However, overall, it appears that the screening with PSC Tool is only completed less than half the time, therefore a gap is present and sufficient for improvement in care. Also, some issues of socioeconomic differences and language (non-English speaking) that may be addressed in future.
- The developers present data indicating wide ranges of screening rates that demonstrate a gap in care as well as disparities by age
- While there is a gap in care, the importance of this gap (screening by this specific instrument) may not reflect a general lack of attention to alternative screening approaches. Disparities are probable but the data are limited as discussed.
- A performance gap was clearly demonstrated by the data and appropriate statistics were provided. There were gaps with age range discrepancies but no significant disparities by race, ethnicity, or language.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: **Specifications** **and** **Testing**

2b. Validity: **Testing**; **Exclusions**; **Risk-Adjustment**; **Meaningful Differences**; **Comparability**; **Missing Data**

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

NQF Staff Evaluation Summary:

Reliability

- Reliability testing conducted at the Patient or Encounter level:
  - Inter-rater reliability was assessed by comparing billing codes indicating that a screen had been given with a review of the progress note to look for evidence of the screen
    - 2021: Chart reviews used as the gold standard. Research assistants examined the WCV progress note and supporting documents to see whether (1) the item and subscale scores downloaded from the EDW were also found in the progress note and (2) whether WCV that did not have item- and subscale-level data for the WCV in the EDW download were also bereft of such data in the patient's EHR.
  - For both the 2017 (using billing codes) and 2021 (using EDW data) data, intercoder agreement was over 90%. Developer states that these levels of agreement indicate an acceptably high level of reliability and therefore both methods of assessing the completion of a PSC screen as a part of a WCV appear to be reliable.

Validity

- Validity testing conducted at the Patient or Encounter level:
  - The developer conducted empirical patient/encounter validity testing comparing the gold standard assessment of the clinical record with reported performance in inter-rater reliability (IRR) to multiple data sources. In the original 2017 submission, developers conducted data element validity testing from three sources in four testing schematics. In 2021, developers gained expanded access to their one of the original data sets and obtained a fourth data set to conduct gold standard patient/encounter (previously data element) validity testing.

        - Data source 1x: Medicaid claims data for WCV from 664 practices/providers with at least 30 WCVs from 1/1/2016 - 12/31/2016 for children enrolled in the state's largest Medicaid HMO, same patients in Data Source 1 which also included all claims data.
            - Testing and results using chart review as the gold standard for validity, the 96110 code in claims data had a sensitivity of 86.2% and a specificity of 100.0% and an overall rate of 91% intercoder agreement in coding the presence of a behavioral health screen during a well child visit. Kappa was .84.
        - Data source 4 (*new data for 2021 testing*): Psychosocial screening and demographic data were extracted from the Electronic Data Warehouse (EDW) for all patients 5.5 to 17.99-years of age seen for WCV from 1/1/2018 - 12/31/2019 from 15 Massachusetts pediatric practice networks covering approximately 100,000 children.
            - Testing and results from two coders reviewed twenty charts selected at random to assess whether the EDW record that a PSC given could be verified in a chart review. Inter-coder reliability was 100%.

- Validity testing conducted at the Accountable level:
    - The developer states face validity was conducted for the measure, however, the information provided does not match NQF's requirements for face validity. The information should be classified as empirical validity testing at the accountable level
        - Developers stated that 304 screening entities in data source 1x were completed and positive screenings differed. The ICC for screening rates (the proportion of total variance in screening rates accounted for by site variation) was 0.38 (95% confidence interval [CI]=0.25- 0.50). With a proportion ≥10.0%, the observed between-site screening rate and a positive screening rate at 245 sites of 0.10 (95% CI=0.00-0.29) can be considered meaningful.
        - The developer also evaluated positive screens and BH service utilization at six months following screens. A positive but statistically non-significant relationship between a site's BH screening rate and the rate of subsequent BH service use was found (Pearson's r=0.08, p=0.17). For testing demonstrated that sites that screened above 71.9% were significantly more likely to have BH treatment rates of 6.0% or above when compared to sites with less than 71.9% screening (88.3% vs. 79.4%, p<.05), suggesting a positive relationship between BH screening rates and rates of subsequent BH service use at the site-level.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- The data elements were clearly defined as was the calculation algorithm. The measure can be consistently implemented.
- Evidence is evaluated based on claims data and reliability is evaluated based on chart review. However, intercoder agreement was 90%.
- Data elements are clear and can be based on chart review or claims data, however, the latter may provide an underestimate of screening. If done for all visits, all ages, all within a population seeking pediatric well care should be equitably screened. Implementation in a busy practice using paper tools will present work effort by office staff to administer, and document results in electronic health records in searchable data elements. Creating digital methods to collect this data will require investments in technology and transfer of data into electronic health record, which can be an expense small practices may see as a challenge. Health literacy may create challenges with either paper or electronic survey methods.
- Inter-rater reliability testing found that intercoder agreement was over 90%. Developer states that these levels of agreement indicate an acceptably high level of reliability and therefore both methods of assessing the completion of a PSC screen as a part of a WCV appear to be reliable
- Clearly these data are reliable in terms of data reporting. The screening is additional service therefore an incentive for the provider. A question as to why missed completely by some sites?
- Measure appears clearly specified, uses claims data.
- The reliability appears to be adequate.
- Elements and calculations are clear. My only concerns regarding consistent implementation of this measure are 1) whether ALL of the most commonly used screening tools will/should be counted and 2) will the insurance/managed care claims data always differentiate between the various instruments. Consistency of applying the screening tool will be provider-variable, with multiple controlling factors in each practice office, such as volume, availability of personnel to administer, etc.

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- The reliability of the measure is moderate based on the data presented. The intercoder reliability was high at 90% suggesting that the data is reproducible. I have no concerns.
- No.
- No concerns, given the performance in a state with mandatory screening, and large health systems that have implemented screening as routine process in well childcare.
- No concerns about the reliability of the measure.
- With all the studies using the PSC Tool, I have little concern about the reliability of the measure.
- No concerns.
- Not my primary concern.
- My concerns regarding reliability of this measure are 1) whether ALL of the most commonly used screening tools will/should be counted and 2) will the insurance/managed care claims data always differentiate between the various instruments. Consistency of applying the screening tool will be variable, with multiple controlling factors in each practice office.

2b1. Validity -Testing: Do you have any concerns with the testing results?

- The developers used the gold standard at the patient or encounter level and face validity at the accountable level. Both were found to be high. I did not have any concerns with the results of the data.
- No concerns. Data source 4, inter-coder reliability was 100%.
- No.
- One study only includes 20 charts. The other study was much larger and showed a sensitivity of 86.2% and a specificity of 100.0% and an overall rate of 91% intercoder agreement in coding the presence of a behavioral health screen during a well child visit. Kappa was .84.
- Although the validity of screening is supported by the data submitted, there are some indications that the validity of the positive screens may need to be better understood with more data collection over time. It seemed there is more to the story that once the screening is done, then the positive screens need better follow up if the health care is to be improved.
- No concerns.
- Yes, I believe validity is quite concerning. The link to even behavioral service utilization, let alone outcomes, is tenuous. I am much more concerned about the lack of a consistent tie to outcomes (see evidence).
- No concerns with testing results.

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- The exclusions were appropriate and consistent with the evidence. There was no risk adjustment.
- No.
- No concerns.
- No concerns.
- Risk adjustment does not appear to be an issue.
- While factors like SES may impact who eventually gets treatment, they should not impact who gets screened.
- I need to delve further on the specs to answer this.
- No concerns about exclusions or risk adjustment.

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- It appears that these analyses are able to measure meaningful differences about quality among the various levels of analysis. They also performed chart reviews to confirm that the expected data was available in the note to confirm the validity of the data. The results are compatible. There was no evidence that missing data was a threat to validity.

- Review should be classified as empirical validity testing at the accountable level, not face validity.
- Health literacy or challenges with use of technology in practices with digital tools to capture may produce missing data for individuals/families with high risk for which screening is beneficial
- Uses code 96110 to identify screening. Does not link screening results to treatment or outcomes. No threats to validity.
- The significant differences in the sites may represent some underlying issues with the patients who utilize the site. They could be threats to validity or not. The measure may be improved with more insights into these issues.
- Results suggest that differences do indicate meaningful differences in quality
- I am not sure that differential rates of using this instrument signify important differences in quality.
- Missing data (providers NOT billing for the screening or billing incorrectly/inappropriately) will affect the numerator.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- Data elements are abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
- Some data elements are in defined fields in electronic sources.
- No proprietary elements are used in implementing this measure. There are no fees, licenses or other requirements needed to use any aspect of the measure or the instrument.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- The data elements are based on claims data or other data routinely used during care delivery. No concerns about feasibility. The developers provided examples of implementation of the measure.
- Claims data
- 1. How will busy practices input data collected on paper into the electronic health record consistently? 2. Can smaller practices afford the digital options for data collection? Literacy as listed in previous responses. In the course of a busy office visit, will a 17 questionnaire be at risk to be "by-passed" due to time constraints of the visit?
- The measure is feasible. Requires abstraction which could increase resource needs to implement measure.
- If PSC screening is done, it appears to be reported via claims.
- Data obtained from claims and in use already
- Certainly, the instrument can be integrated into practice, an EHR and an overall quality process. But such data are not routinely captured.
- No concerns. If the provider screens and bills for the screen, then it will show up in the claims data - if ALL CLAIMS data is requested and not just PAID CLAIMS.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

Accountability program details

- This measure is used for Professional Certification or Recognition Program: American Board of Pediatrics MOC credit for all Affiliated Pediatric Practices (APP) network providers, Reliant Medical Group (RMG) MOC project.
- This measure is used for Quality improvement with Benchmarking (external benchmarking to multiple organizations).
- The Children’s' Behavioral Health Initiative (CBHI) requires that all pediatric providers in the Commonwealth of Massachusetts who bill for a WCV must have used one of about twelve approved measures to screen for psychosocial problems as a part of that WCV.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured, and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Psychosocial screening reports which include item and subscale data, and interpretation of risk scores are available to clinicians and patients instantaneously in the EHR.

Additional Feedback: N/A

Questions for the Committee:

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluates the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The developer highlighted the increase in rates (39.98% from 12/31/07-12/31/2008 to 70.69% from 1/1/2017-3/31/2017) of screening in the full sample of 3–17-year-olds over the first 9.25 years of CBHI.
- The developer noted that there is an increasing number of networks that have adopted routine psychosocial screening as the standard of care.
- The developer also noted that the country of Peru has made routine psychosocial screening a required part of its annual pediatric care for all adolescents and has specified the PSC as the measure that must be used.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer notes the following unintended benefit of implementing this measure:
    - Increasing widespread use of a simple but effective PRO tool that can be used for screening, diagnosis, and the monitoring of treatment outcomes for psychosocial problems (Massachusetts, California, and other states).
    - Increased national use of the measure (PSC is being used in the SAMHSA National System of Care Expansion Evaluation and in the state of California child mental health outcomes assessment program), Mental Health America making the PSC and PSC-Y available for free and to tens of thousands of youths or their parents.
    - Increased use of the PSC as a mandatory tool in behavioral health networks as a way to meet the QI goal of using standardized measures in outcomes-based care. The state of

California's requirement that the PSC be administered at intake for mental health services and then every six months and the Community Behavioral Health Association of Maryland's similar requirement in its 72 member agencies are examples of this.

Potential harms

- No unintended negative consequences to individuals or populations.

Additional Feedback:

N/A

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided?4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- The data is in use for quality certifications. The results of the screening data is immediately available for clinical care decisions.
- Pass
- Public reporting via required screening by some states. Data can be used for maintenance of certification for American Board of Pediatrics. Use is easy, if time allows and parents/children are willing, and time pre-visit allow. Scoring is easy, and threshold scores are provided for when to consider intervention.
- PSC is being used in the SAMHSA National System of Care Expansion Evaluation and in the state of California child mental health outcomes assessment program. Also being mandated by the country of Peru.
- The PSC appears to be widely used and accepted as a standard (AAP and others) with revisions as feedback is incorporated. The inclusion of the screening tool is easily evaluated by the claims data. Use is Pass.
- Yes - The developer highlighted the increase in rates (39.98% from 12/31/07-12/31/2008 to 70.69% from 1/1/2017-3/31/2017) of screening in the full sample of 3–17-year-olds over the first 9.25 years of CBHI.
- It is being used, but not quite in the way this measure would be used as a stand-alone, i.e.., Mass. uses as one of a number of approved measures and in a quality improvement mode.

- I agree with the developers answers to 4a1 and 4a2. Speaking for our very large, national MCD, CHIP, and Exchange MCO, we do give performance results and data, as well as assistance with interpreting the measure results, and give an opportunity to provide feedback on their measure performance.

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- There is evidence of performance improvement over time although a slight downturn during COVID. There do not appear to be potential harms associated with this measure since using such a screening tool can improve clinical outcomes.
- Like with all screening tools, they are most effective when the individuals conducting the screening have strong linkages to the referral resources that patients might require for follow up care. Nationally, access to children's behavioral health services is still limited. Also, patients are more likely to follow up with MH care when MH services are co-located and integrated into the practice. Widespread adoption of integrated primary care and behavioral health models for pediatrics is still not the norm. So, implementing the screening tool requires that pediatricians have more specialized training and or supports to provide brief counseling and referral to care.
- Usability is good, and the developer provides data on steady increased screening rates. Benefits include early identification of those with needs that can be addresses early. Unintended consequence that I do not see mentioned is potential for questionnaire fatigue.
- The results could increase treatment for screening identified needs. No identified harms.
- Since the measure provides accountability and transparency, its benefits outweigh the harms as presented. It would be interesting to know the subscale changes over time as the pandemic may be causing some greater issues in behavioral health.
- No harms identified.
- There really wasn't a substantive discussion of harms--e.g., does screening place added burden on the health system by identifying many cases of little clinical impact; is screening itself a good or is it the tie to improved outcomes? As in many areas of behavioral health, we have learned that screening, without effective follow up, and clear demonstration of impact on outcomes, may have many underestimated harms. I don't concur with the staff's evaluation of high.
- Success with the measure will further quality care and increase the efficiency of the system by ensuring that screening occurs routinely and as part of the larger well-child visit, killing two birds with one stone. No unintended consequences.

Criterion 5: Related and Competing Measures

Related or competing measures

- 2801: Use of First-Line Psychosocial Care for Children and Adolescents on Antipsychotics
- 1365: Child and Adolescent Major Depressive Disorder (MDD): Suicide Risk Assessment
- 0576: Follow-Up After Hospitalization for Mental Illness (FUH)
- 0108: Follow-Up Care for Children Prescribed ADHD Medication (ADD)
- 0711: Depression Remission at Six Months

- 1885: Depression Response at Twelve Months- Progress Towards Remission
- 0712: Depression Assessment with PHQ-9/ PHQ-9M
- 1884: Depression Response at Six Months- Progress Towards Remission
- 0710: Depression Remission at Twelve Months

Harmonization

- Out of the nine related measures identified, four do not overlap with the PSC (#0108, #0576, #1365, and #2801).
- The remaining five NQF-endorsed pediatric mental health measures relate to the PSC because all involve depression and rely on the Patient Health Questionnaire (PHQ)-9.
- The PSC does not compete with these five adolescent depression measures because the PSC does not have the same target population and has a much broader focus.
- All five of the currently endorsed measures that use the PHQ-9 apply only to adolescents already diagnosed with depression; in contrast, the target population for the PSC is children as well as adolescents and it includes 100% of both children and adolescents, not just the 5-10% of adolescents who are depressed.
- The developer indicated that there are two non-NQF endorsed instruments that are related to the PSC, namely, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Child Behavior Checklist (CBCL; Achenbach, 1991).

Committee Pre-evaluation Comments: Criterion 5:

Related and Competing Measures

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- There are several related or competing measures but there does not appear to be overlap that needs further harmonization.
- 9 similar measures but most are for individuals already identified as having a BH diagnosis.
- No need for harmonization of 5 overlapping measures. However, the use of the PSC and a depression measure in those adolescents being evaluated and treated for depression may add survey burden to these adolescents during visits.
- Five NQF endorsed measures compete with this measure. This measure includes all children (and adolescents). Competing measures have a narrower eligible population.
- There are two non-NQF measures, but unclear how they are used for screening or other measures. Also are the sub-scales similar?
- No real overlap since other measures involve populations with an existing diagnosis
- My major concern is the use of a single instrument of moderate complexity and burden. The other measures are only tangentially related.
- Not to my knowledge. It appears that all NQF-endorsed related measures are NOT regarding routine BH/MH screening in "well children," not already identified with an issue.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/12/22

- **No NQF Members have submitted support/non-support choices as of this date.**
- **No Public or NQF Member comments submitted as of this date.**

Staff Scientific Acceptability Evaluation

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "Measure_3332_MIF" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

RELIABILITY: TESTING

Submission document: "Measure_3332_MIFMIF" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☐ **Measure score** ☒ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical *VALIDITY* testing** of **patient-level data** conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and **all** testing results):

    ☐ **High** (NOTE: Can be HIGH **only if** score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has **not** been conducted)

    ☐ **Low** (NOTE: Should rate **LOW** if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate **INSUFFICIENT** if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

- The specifications are precise, unambiguous, and complete for consistent implementation (Box 1) >>> Empirical reliability testing was conducted (Box 2) >>> Accountable level reliability testing was not conducted (Box 4) >>> Empirical patient/encounter reliability testing was conducted (Box 8) >>> Method was appropriate (box 9) >>> Moderate

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

- The developer reports five methods for identifying meaningful differences in performance from the submitted data sources, literature, and validity testing. No concerns were identified.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.

- The developer reported testing from multiple sources to conduct testing. No specific concerns were identified.

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.

- The developer states they did not assess missing data and state that "Missing data, in this case PSC screening scores is not an issue since those patients who have WCV and are not assessed in the measurement period remain in the denominator."

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☒ Yes ☐ No ☐ Not applicable

    16c. **Social risk adjustment:** NA. The process measure is not risk adjusted**.**

    16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

16d.**Risk adjustment summary:** NA. The process measure is not risk adjusted**.**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No
16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No
16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No
16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No
16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

For cost/resource use measures ONLY:

17. **Are the specifications in alignment with the stated measure intent?**

☐ Yes ☐ Somewhat ☐ No (If "Somewhat" or "No", please explain) NA. This is not a cost/resource use measure.

Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers): NA. This is not a cost/resource use measure.

VALIDITY: TESTING

18. **Validity testing level:** ☐ **Measure score** ☐ **Data element** ☒ **Both**

19. **Method of establishing validity of the measure score:**

☐ **Face validity**

☒ **Empirical validity testing of the measure score**

☐ **N/A (score-level testing not conducted)**

20. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

21. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

22. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

23. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

Submission document***:*** *Testing attachment, section 2b1.*

☒ **Yes**

☐ **No**

☐ **Not applicable** (data element testing was not performed)

24. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there **are** threats to validity and/or relevant threats to validity were **not assessed OR** if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level **is required**; if not conducted, should rate as INSUFFICIENT.)

25. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

- All threats to validity were adequately assessed (Box 1) >>> Empirical validity testing was conducted using the measure as specified (Box 2) >>> Validity testing was conducted of the accountable entities using multiple statistical methods (Box 5) >>> Conducted validity methods were described and appropriate for assessing conceptual and theoretically hypothesized relationships (Box 6) <<< Based on the validity results and the completeness of the patient/encounter and accountable entity level testing (highest possible rate is high) (Box 7a) >>> Moderate

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

26. **What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?**

☐ **High**

☐ **Moderate**

☐ **Low**

☐ **Insufficient**

☒ **NA. This is not a composite measure.**

27. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**

NA. This is not a composite measure.

ADDITIONAL RECOMMENDATIONS

28. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

No additional concerns were identified.

Developer Submission

1. Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria

---

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

[Response Begins]
Yes

2021 Submission:

Since our last NQF submission in 2017, there has been new evidence about the measure, both in the form of item level data (which we have labeled Data Source 1x for a large subsample from our original Data Source 1) and an entirely new sample (from a large network of pediatric practices serving predominantly commercially insured outpatients who were screened exclusively with the PSC using a fully electronic system (Data Source 4).

In this and another pediatric network, routine psychosocial screening with the PSC was the basis of a Maintenance of Certification program for pediatricians.

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

2021 Submission:
Updated evidence information here.

2018 Submission:
Evidence from the previous submission here.

Evidence

1a.01. Provide a logic model.

Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

[Response Begins]

2017 Submission
Below, we list a logic model for the completion of the PSC tool as a required component of a pediatric WCV. This process measure increases the likelihood that psychosocial issues have been identified, explored, and dealt with during the WCV. There is now strong evidence that routine screening with the PSC Tool leads to a significant increase in the number of children with problems who receive outpatient mental health treatment. There is also moderate to strong evidence that children who are screened and receive services in this way show significant reductions in symptoms and improved short- and long-term mental health and functional outcomes.

2021 Submission:
Step 1) Screen:
Screen all children ages 4-17 with the Pediatric Symptom Checklist at a well-child visit.
Step 2) Assess the Outcome and Prompt a Conversation:
The provider reviews the score on the PSC. If the child scored "At risk" the provider will discuss any concerns and the option of additional information or services.

Step 3) Refer for Additional Services:

Parent and provider decide on next step: watchful waiting, counseling by pediatrician, referral to specialty mental health provider, or other service.

Step 4) Receive Mental Health Services:

A higher percentage of children with psychosocial problems receive mental health services which in turn leads to fewer psychiatric symptoms and better functioning at school, home, and with friends.

[Response Ends]

1a.02. Select the type of source for the systematic review of the body of evidence that supports the performance measure.

A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data.
[Response Begins]
Other (specify)
2021 Submission:
Long-established clinical practice and expert opinion from the Task Force on Mental Health of the American Academy of Pediatrics, as documented in an evidence review (AAP, 2010).
[Response Ends]

If the evidence is not based on a systematic review, skip to the end of the section, and do not complete the repeatable question group below. If you wish to include more than one systematic review, add additional tables by clicking “Add” after the final question in the group.
Evidence - Systematic Reviews Table (Repeatable)
Group 1 - Evidence - Systematic Reviews Table

1a.03. Provide the title, author, date, citation (including page number) and URL for the systematic review.

[Response Begins]

N/A
[Response Ends]

1a.04. Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

[Response Begins]
N/A
[Response Ends]

1a.05. Provide the grade assigned to the evidence associated with the recommendation and include the definition of the grade.

[Response Begins]
N/A
[Response Ends]

1a.06. Provide all other grades and definitions from the evidence grading system.

[Response Begins]
N/A
[Response Ends]

1a.07. Provide the grade assigned to the recommendation, with definition of the grade.

[Response Begins]
N/A
[Response Ends]

1a.08. Provide all other grades and definitions from the recommendation grading system.

[Response Begins]
N/A
[Response Ends]

1a.09. Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

[Response Begins]
N/A
[Response Ends]

1a.10. Provide the estimates of benefit, and consistency across studies.

[Response Begins]
N/A
[Response Ends]

1a.11. Indicate what, if any, harms were identified in the study.

[Response Begins]
N/A
[Response Ends]

1a.12. Identify any new studies conducted since the systematic review and indicate whether the new studies change the conclusions from the systematic review.

[Response Begins]
N/A
[Response Ends]

1a.13. If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, describe the evidence on which you are basing the performance measure.

[Response Begins]
2021 Submission:
Although not the subject of a rigorous systematic review or official clinical practice guideline, screening for psychosocial problems as a routine part of pediatric well child visits has been a recommended practice by the American Academy of Pediatrics since the publication of *Bright Futures* two decades ago. In their "Recommendations for Preventive Pediatric Health Care" (https://downloads.aap.org/AAP/PDF/periodicity_schedule.pdf) Periodicity Schedule, the American Academy of Pediatrics and Bright Futures specify the inclusion of "psychosocial/behavioral assessment" at every well child visit from birth to age 21. These recommendations are based at least in part on "The Case for Routine Mental Health Screening" that was made by the American Academy of Pediatrics Task Force on Mental Health (AAP 2010) with these recommendations recently officially reaffirmed (AAP, 2021).
[Response Ends]

1a.14. Briefly synthesize the evidence that supports the measure.

[Response Begins]
2021 Submission
For school aged children (the age range of the PSC), the AAP's recommendation is to "Use validated instruments to screen all school-aged children (5 years through adolescence) for symptoms of mental illness and impaired psychosocial functioning at health maintenance visits; at any time of family disruption, poor school performance, reported behavioral difficulties, recurrent somatic complaints, or involvement of a social service or juvenile justice agency; and when child or family identifies psychosocial concerns." This recommendation was made by the AAP (2010) on the basis of the following evidence:

"A number of mental health disorders, in addition to depression, are prevalent in children and adolescents, e.g., behavior or conduct problems and attention-deficit/hyperactivity disorder in 6.3% and 8.8%, respectively, of children 6 to 17 years old; anxiety disorders in 16% of children 9 to 17 years old. There are no systematic reviews of clinical outcomes related to isolated screening for these individual conditions or for general psychosocial screening of children and adolescents. However, prevention and early intervention efforts targeted to children, youth, and families have been shown to be cost-effective, reducing use of more costly services such as welfare dependency and juvenile detention. Emotional and behavioral problems in young children may persist or worsen and adversely affect early and later school performance, and children from poor families are generally at greater risk. These findings suggest that early detection and intervention, particularly in low-income populations, may prevent or ameliorate mental health problems in children and adolescents." (AAP 2010).

As noted above, the AAP also endorses routine screening in its periodicity schedule for 'Preventive pediatric health care' that recommends, "an assessment of emotional and behavioral health" at every well child visit. The citation for this specific recommendation in the schedule ~~chart~~ is a 2015 paper entitled 'Promoting Optimal Development: Screening for Behavioral and Emotional Problems' Weitzman, et al, 2015). This paper was co-authored by the AAP Section on Developmental and Behavioral Pediatrics, the AAP Committee on Psychosocial Aspects of Child and Family Health, the AAP Council on Early Childhood, and the Society for Developmental and Behavioral Pediatrics. The paper supports its recommendations by citing 69 references that review studies on the prevalence of behavioral and emotional disorders, the continuing lack of treatment for them, factors affecting prevalence and screening, and recommended screening measures. The PSC is one of just a handful of measures mentioned in this paper for broadband screening of school aged children and adolescents. The AAP has also recently (Foy et al., 2019) updated its 2009 Policy Statement on Mental Health Competencies for Pediatric Practice (Committee on Psychosocial Aspects of Child Family Health & Task Force on Mental Health, 2009), continuing to stress the importance of using mental health screening tools as one key mental health competency that pediatricians should have.

In this section of the MIMS form we will briefly synthesize the evidence from papers published in the past four years that support the measure and have made 'the case for routine psychosocial screening' in pediatrics even stronger during that time. One of the most important pieces of evidence for any proposed measure is the endorsement of relevant professional organizations and the American Academy of Pediatrics is arguably the most important for a pediatric measure. The fact that **in May of 2021,the AAP officially reaffirmed its 2015 position, and continues to recommend an assessment of emotional and behavioral health as a part of every well child visit for every child of every age (AAP, 2021*), providing strong support for NQF's continued endorsement of psychosocial screening using the Pediatric Symptom Checklist Tool.***

Another type of evidence includes empirical and review studies. Since our 2017 NQF endorsement proposal, 120 papers have been published that either used the PSC as a measure or reviewed it along with other measures. The citations for these 120 papers are listed in Appendix 1 to this proposal, grouped under a half dozen general categories that capture the papers' main focus (e.g., "Screening with the PSC is now the standard of care in many real-world settings" and "Additional PSC projects are demonstrating the feasibility and outcomes of screening," etc.). Appendix 2 adds the reference list for these, and all of the other studies cited in this 2021 re-endorsement proposal including, for the convenience of the reader, all of the older papers that we referenced in our 2017 NQF endorsement proposal.

To summarize this many papers in just a few pages, we will present the studies as they relate to the three points listed in the AAP's 2010 original 'case for routine mental health screening': *"(a) Does use of a validated tool accurately identify children with mental health problems or improve identification of children with psychosocial problems? (b) Do the identification of problems and linkages to services improve outcomes? (c) What is the feasibility of screening in a busy primary care practice?"* We will summarize what recent PSC papers have said about these questions in reverse alphabetical order, beginning with question "c" which asks about the feasibility of screening in the real world.

Since the submission of our 2017 NQF proposal, more than a dozen newly published PSC papers have demonstrated the feasibility of screening by documenting the fact that **screening with the PSC is now the standard of care in many real-world settings and/or that it has been evaluated in real world demonstration projects.**

The strongest evidence for the feasibility of PSC screening in the real world that was presented in our 2017 NQF proposal was the Commonwealth of Massachusetts Children's Behavioral Health Initiative (Kuhlthau et al., 2011; Romano-Clarke et al., 2014). As our 2017 proposal noted, beginning in 2008, Massachusetts required that a brief psychosocial screen be administered at every well child visit of every pediatric patient with Medicaid. A state website (https://www.mass.gov/info-details/cbhi-data-reports) shows the number of screens for each quarter of each year since the start of CBHI and in the past four years, more than a million screens have been administered (with an estimated half million of them PSCs). The fact that the CBHI program is about to enter its 15th year and that it is still going strong is still the strongest evidence of the feasibility of such a program. Additional evidence of feasibility is the fact that even though the rate of psychosocial screening (and of WCVs) dipped for a few months at the beginning of the Covid-19 pandemic, it appears to have returned to its previous level. It is important to note that the CBHI program continues despite the ending of the consent decree (Rosie D vs Romney; see Kuhlthau et al., 2011) that led to the screening mandate in the first place.

Over the past four years, there have been more than a dozen new studies that have also provided evidence of the feasibility of psychosocial screening. The studies are listed under the section entitled “I.A. Screening with the PSC is now the standard of care in many real-world settings” in Appendix 1. Two of the most relevant of these studies have been conducted as a part of routine care in pediatric networks in which the majority of patients are insured with commercial rather than Medicaid insurance (in most previous studies it has been the opposite). Murphy and colleagues (2021) documented the rollout of a fully electronic approach to screening in a network of 18 suburban pediatric practices. The network had adopted screening with the PSC as a best practice for all WCVs with children aged 5.5 to 17 years old. More than 75,000 patients were seen for well child visits over a two-year period and the investigators reported that nearly 90% of the patients were actually screened in the first year and that even in the second year, nearly 80% of all WCVs had completed screens.

Parenthetically, it may also be important to note that although the above paper reported only on screening patients who were 5.5-17.9 years of age, an additional 30,000+ patients aged 0 to 5.5 years old were also successfully screened with the Baby and Preschool versions of the PSC as a part of the same program. Arauz-Boudreau and her colleagues (2020) described a similarly successful electronically based PSC screening program in a single clinic that was subsequently implemented across a seven-clinic network. And Holcomb et al. (2021) reported on how in the same system--with the PSC given at all WCV--the PSC internalizing scale could be used to track problems and referrals for depression through repeat screenings over a one-year interval. Further evidence of the feasibility of routine screening is the fact that in both of these large real-world programs, screening with the PSC continues to be the standard of care even after several years.

Shellman and colleagues (2019) reported on the results of a three-year rollout of a protocol for using the PSC-17 as a routine screen in 3,678 WCV in a single clinic affiliated with a large urban children’s hospital system in Texas. The success of this program led to its adoption as the standard of care across the larger integrated delivery system of 52 affiliated pediatric primary care practices. Spencer and her colleagues (2020) have recently reported on the relationship between PSC-17 scores and scores on a parent survey of social determinants of health, both of which have been required components of WCVs in a hospital-based pediatric clinic for more than five years. McLaurin-Jiang and associates (2020) presented data from a single site pediatric continuity clinic which has about 19,000 WCVs per year and which has administered the PSC-17 and PSC-Y as the standard of care since 2011. This study explored the impact of a program that taught pediatric residents about an integrated mental health approach, finding that resident confidence and the use of secondary screens increased significantly.

Burke and colleagues (2021) reported that it was feasible to implement a screening and referral initiative using the PSC-17-Y in the school-based health center of one New York City middle school. Of the 741 students screened, more than one quarter were found to be at risk and more than eighty percent of them accepted a referral to either a mental health or a pediatric clinician. Hansel and her associates (2017) showed that routine screening with the PSC-17 was feasible in a demonstration project at two Federally Qualified Health Centers in rural Louisiana.

Honigfeld et al. (2017) reported on a Connecticut project involving more than 1,200 patients from seven suburban pediatric practices whose parents filled out the PSC at their WCVs. The authors found that using an electronic system that added a mental health history questionnaire and presented best practices and referral options to the PSC increased the success of screening, including higher rates of follow-through for next level assessments. Young and Takala (2018) explored the possibility of using two screens sequentially (either the PSC, the SDQ, or the CBCL), using simulation modeling to show that administering two screens in sequence could lead to improved accuracy.

Jacobson et al. (2019) described the feasibility and correlates of having foster parents fill out the PSC-17 in a sample that included all 6,492 children in out of home care in Washington state from 2010 to 2015, concluding that it could be useful as an outcome and tracking measure (as well as an intake measure) for that state in the future. And although there have been no published studies, a website for the state of Minnesota documents the fact that since 2003, the state of Minnesota has used the PSC as a required screen for all school aged children entering or continuing in foster care each year (https://mn.gov/dhs/partners-and-providers/policies-procedures/childrens-mental-health/screening/).

Section I. B in Appendix 1 lists a half dozen additional PSC projects that have shown the ~~demonstrated~~ feasibility and outcomes of screening in **demonstration projects**. Berger-Jenkins and colleagues (2019) reported on one such demonstration that implemented psychosocial screening in a busy urban pediatric practice over an 18-month period that ultimately screened more than 2000 children using the Baby and Preschool PSC as well as the parent PSC-17 and a measure of social determinants of health. Screening rates reached 90% and most children who screened positive were seen for follow up. Stadnick and colleagues (2020) reported on a mixed methods study that explored the opinions of sixty pediatric primary care providers and leaders from three organizations about providing integrated mental health care for children with autism. Mental health screening in primary care was one of the most important strategies and the PSC was one of the most frequently used measures. Omkarappa and colleagues (2021) used the PSC-Y to screen 480 students aged 12-16 from six high schools in India to provide a high-risk sample that then documented the positive impact of a group cognitive-behavioral therapy for internalizing behavior problems among children of parents with alcohol use disorders.

In another study based in India, Muppidathi (2017) and colleagues conducted a study of the prevalence of problems found in 450 patients whose parents completed the PSC-17 in a hospital-based outpatient pediatric clinic in India. The authors noted that the prevalence of risk scores on the PSC and the impact of sociodemographic variables on PSC risk rates were comparable to those reported in US studies, leading the authors to conclude that “the consistency and widespread acceptability of the PSC make it well suited for the next generation of pediatric mental health services research, which can address whether earlier recognition and intervention for psychosocial problems in pediatrics will lead to cost effective outcomes.” This statement could be viewed as one of the most succinct summaries of the current status of research on psychosocial screening with the PSC…fully validated as accurate and feasible in a wide range of real-world settings and ready to be used in more definitive studies of population level outcomes.

One of the largest PSC screening programs (and strong evidence for the feasibility of routine psychosocial screening) comes from Chile (Guzmán et al., 2015). As noted in our 2017 NQF proposal, for more than twenty years, the national Ministry of Education in Chile has used the PSC and one other measure (the Teacher Observation of Classroom Adaptation or TOCA) to screen all students in what is now more than 2000 participating schools each year. In this program, called Skills for Life (SFL), positively screened students are referred to in-school preventive workshops and there is evidence that those who attend show significant improvements in both PSC and TOCA scores and academic functioning (Guzmán et al., 2015). According to a recent paper (Leiva et al., 2021) more than 150,000 first and third grade students were screened in 2018, as were tens of thousands of other students from the middle- and high school levels of the program. Since the program’s inception in 1999, SFL reports that more than one million students have been screened, most of them with the PSC and most longitudinally from first to third, or sixth through eighth grades.

Further evidence of feasibility based on large numbers can be seen at the start of Appendix 1 to this proposal, which begins with a figure that graphically illustrates the number of PSC papers published each year from the first in 1979 to the present (October 21, 2021). The figure is important because it illustrates the accelerating increase in the number of PSC publications, growing from an average of about one publication per year for its first decade, to three a year over its second decade, to twenty to thirty publications per year over the past decade and one half. This continuing and growing use of the PSC can be viewed as additional evidence of its feasibility.

Supplementing these published accounts of real-world screening are anecdotal reports which, although not academically confirmed, are important in portraying an even broader picture of real-world feasibility. When the

Covid-19 pandemic struck in 2020, all schools in Chile, as in most of the rest of the world, switched to remote learning and, since it could be administered electronically over cell phones and computers (the TOCA cannot since it is based on interviews with teachers), the PSC was the only screen administered by the SFL program in 2020 and 2021. Despite the many logistical difficulties caused by the pandemic, more than 70,000 PSCs were completed in 2020 and although the data are not yet available, program officials believe that even more completed PSC screens were collected in 2021 (Simonson, 2021, personal communication). The government of Peru has recently officially adopted the PSC as the primary screen for mental health in its national pediatric health program with more than ten thousand PSC-Y forms already completed (Cabellero, personal communication).

In what may be the largest screening program in the US, a company called CHADIS provides pediatricians with software that administers the PSC and other measures electronically before well child visits. A paper cited in our previous NQF proposal (Murphy et al., 2016) was based on a national sample of more than 100,000 PSC-17s administered by CHADIS over a decade ending in 2015. The rate of screening has increased geometrically since that time, with more than 750,000 PSCs administered from the beginning of 2019 through mid-October 2021 (Sturner, personal communication). As noted earlier about the number of PSC papers, the accelerating increase in PSC screens as well as the large number of them, provide important evidence of the feasibility of routine screening with the PSC. Further evidence of the feasibility of PSC screening is the large number of PSC's administered by Mental Health America. As described in a 2018 paper (Murphy et al., 2018), MHA offers the PSC and a handful of other brief screens on its website, providing free score reports to anyone who wants one and giving adolescents and adults the opportunity to self-screen. Since the beginning of 2019, more than 500,000 PSC's and PSC-Y's have been completed through the MHA website (Nguyen, personal communication).

It is also important to add that real-world screening programs using the PSC in pediatric practices appear to be growing more common in recent years. Although there are no publications or precise data, we do know through studies in progress that we ourselves are doing that the 14-practice Reliant Medical Group now screens all school-aged children and adolescents with the PSC-17 (Dalal, personal communication) each year, as does the seven-site Pentucket Medical Group (Maddox, personal communication), and networks of outpatient pediatric sites affiliated with Boston Medical Center (Adams, personal communication), Children's Hospital Boston (Woodberry, personal communication), and Children's Hospital of Philadelphia (Young, personal communication). We have also heard that Duke, Mayo Clinic, Nationwide Children's Hospital, and San Diego Children's Hospital have implemented electronic screening with the full age range of PSC forms in their networks. As one recent paper (Honigfeld et al., 2017) observed, the increase in pediatric psychosocial screening has probably been aided by the fact that screens like the PSC are reimbursable by Medicaid and many commercial insurers, and that electronic medical record systems like Epic now have patient portals that make it relatively easy for pediatricians to send links requesting that parents fill out a PSC or other Patient Reported Outcome Measures prior to WCV.

Returning to the AAP's 2010 paper that made 'the case for routine psychosocial screening', the next question we will address is "b", whether "*the identification of problems and linkages to services improve outcomes"*? Although, as we will discuss below, many of the papers cited in the previous section on feasibility provide data on outcomes, as noted before, the most important program may still be the Massachusetts CBHI. We will summarize its main findings about outcomes briefly before moving on to more recent studies.

As discussed in our 2017 proposal, in the mid 2010's Hacker and her associates published a series of studies (2014a; 2016; 2014b) based on data from the first few years of CBHI. Using sample sizes ranging from hundreds of thousands (all of the WCV for all children with Medicaid in Massachusetts) to millions of cases (all of the WCV from MA in 2008 and all of the WCV from California in the same year), these studies showed definitively not only that the Massachusetts pediatric psychosocial screening program (in which most of the school aged children had been screened with the PSC) was feasible (70%+ of WCVs had completed screens) **but also that there had been substantial increases in the number and percentages of children with problems who were identified and who then went on to receive outpatient mental health services**. In our view, these indicators are still the most reasonable outcome measures we have for now and the huge sample sizes make it seem likely that the findings will hold up in future studies.

As noted above, papers using data from the CBHI continue to be published. In the time since our last NQF proposal, Murphy and colleagues (2020) reviewed billing records for a more recent year of the CBHI program for all 76,752 children in the state's largest HMO and found significant differences in rates of screening and of rates of positive screening between clinics as well as some evidence that in clinics with higher rates of screening, children who had screened as at risk were more likely to receive outpatient mental health services. The study replicates the findings about increased mental health services for positive screened children from the Hacker et al. studies as well as providing first time ever data that shows significant differences by site on a site-by-site basis. Being able to document significant difference by site is an important criterion for quality measures.

Our 2017 proposal also relied heavily on a series of studies by Kolko and his colleagues (2014; 2010; 2011) to make the case for positive outcomes related to screening. Using methodologically rigorous RCT designs, these studies had provided experimental evidence that children who had screened positive on the externalizing subscale of the PSC were more likely to receive mental health services and more likely to show improved outcomes like lowered scores on measures of behavior problems, hyperactivity, and internalizing problems. Over the past few years, these researchers have published two additional experimental studies of a collaborative care approach to treatment in pediatrics. One study (Kolko et al., 2020) showed significantly improved outcomes for ADHD, oppositional defiant disorder, and quality of life and the other (Shaffer et al., 2017) showed decreased negative parenting for children who had been screened at-risk on the PSC-17 externalizing scale who had received a brief intervention.

Over the past four years, a number of the new studies of real-world programs that were discussed in the previous section as evidence of feasibility have also reported improved outcomes for children identified with psychosocial problems on the PSC. The study of 177 pediatric outpatients by Hansel and colleagues (2017) in two Louisiana health centers reported that for children who screened positive on the PSC, those scores and parenting stress scores decreased and parent satisfaction with care increased in the sample after a brief intervention. The Shellman et al. (2019) study found that a PSC based universal psychosocial screening protocol was cost effective and associated with higher rates of behavioral health treatment for positively screened children. Burke and colleagues (2021) reported that most of the youth who screened positive on the PSC had enhanced access to and connection with mental health services. Berger-Jenkins (2019) reported that among children who screened positive on the PSC or a measure of social stressors, 80% followed up with their medical doctors and 50% completed referrals to clinic social workers.

In another real-world PSC screening program, Holcomb et al. (2021) compared scores longitudinally over two annual WCVs and found that children who screened positive on the PSC-INT scale were 9 times more likely to receive subsequent BH services than were children who screened negative on the PSC-INT. In the paper by Arauz-Boudreau and colleagues (2019) describing overall PSC scores in the same real-world sample, a review of progress notes for a sample of the WCV visits showed that in about 90% of all cases there was written evidence that the pediatrician had discussed the PSC score with the parent. Omkarappa (2021) reported that children of alcoholic parents who screened positive on the PSC and who were randomly assigned to CBT showed significantly improved scores on measures of depression, anxiety, and self-esteem than similar children who did not receive the intervention.

Another group of studies, listed in Section C in Appendix 1, included the PSC as a measure in QI studies in real-world settings. Craig and colleagues (2018) implemented a small QI project in an attempt to raise PSC-17 screening rates in a pediatric resident clinic with a largely underserved patient population, where barriers to screening included parents not having time to complete the form, residents not having time to review and document the screenings, and inadequate referrals. After the QI implementation, screening completion rates increased from 22% to 82%, documentation increased from 38% to 76%, and referrals for patients screened at-risk reached 100%. In another study, Riobueno-Naylor and colleagues (2019) evaluated an electronic system that provided enhanced psychosocial information (scores from the PSC-17 and the Burn Outcomes Questionnaire) to burn care clinicians before appointments so that they could incorporate more of this kind of information into their outpatient care. Both clinicians and parents rated the additional information as very valuable. May and colleagues (2018) outlined an ambitious comprehensive model for providing more integrated care for children with Medicaid. Routine screening with the PSC-17 to identify children's mental health problems is a key component of this plan.

Returning to the AAP's 2010 'case for routine mental health screening', the third question was whether "*the use of a validated tool accurately identifies children with mental health problems or improve[s] identification of children with psychosocial problems?* The majority of the 120 PSC studies done since 2017 provide information that is relevant to this question. Because there are so many, we have broken them down further as they relate to three general topics that are used to structure Appendix 1.

Studies of reliability and validity (see Appendix 1, Section I. D)

The first category covers studies of the reliability, validity, or factor structure of one of the PSC forms (e.g., the 17- or 35-item versions, parent and youth report versions, translations into other languages) and other papers explored related topics. For example, Chaffin and colleagues (2017) added items about functional impairment and behavioral health services to the PSC-17P and reported that the combination form had improved classification accuracy using the Child Behavior Checklist (CBCL) as the gold standard in a sample of 267 youth ages 6 through 16. Bergmann and colleagues (2020) found evidence for PSC-17Y's validity in a sample of more than 19,000 11- to 17-year-olds who completed screenings via the Mental Health America website. Parker and colleagues (2019) reported that both self-reported and foster parent-reported PSC-17s were valid in a sample of 2,389 youth aged 11 through 17 in the Washington state foster care system.

DiStefano and colleagues (2017) found support for the three-factor structure of the PSC-17 using a teacher-rated version of the form in a sample of 836 preschool students in the U.S. Using the same preschool sample, Liu and colleagues (2020) reported that the teacher-reported PSC-17 exhibited gender invariance and other features of statistical validity. In another sample, Liu and colleagues (2020) also confirmed the factor structure of a teacher-reported PSC-17 as well as its measurement invariance (across gender, ethnicity, and grade level) in a sample of 508 elementary school students. Likewise, Studts and colleagues (2017) tested four PSC-17P items for measurement bias across child race, sex, and socioeconomic status in 900 preschool children in primary care and found that three of the four items were free of bias. In another study, DiStefano and colleagues (2019) reported that the positioning of PSC-17 items (i.e., grouped by subscale or mixed across subscales) had no effect on screening results.

Öner and colleagues (2019) established the validity and reliability of the pictorial PSC in 799 Turkish 4- to 5-year-olds. More recently, Ardıç and colleagues (2020) showed that these findings extended to 6- through 16-year-olds using a sample of 729 Turkish students. Chaurasiya and colleagues (2019) reported on the reliability and validity of a Hindi version of the PSC-35Y in a sample of 300 elementary school students. Leiva and colleagues (2019) confirmed the three-factor structure of a Spanish version of the PSC-17P in 5,177 Chilean first graders. Similarly, Higuchi and associates (2021) confirmed the factor structure and reliability of the Japanese PSC-17Y in a sample of 217 students in 5th-6th grades and 84 students in 7th-9th grades.

The PSC as a proxy for mental health in studies of chronic illness and other situations (see Appendix 1, Section I.E)

Many of the PSC papers published in the past four years have relevance for the AAP 2010 paper's question about the accuracy of the screen: *"whether the use of a validated tool accurately identifies children with mental health problems or improves identification of children with psychosocial problems."* Over the years, one of the most common uses of the PSC has been as a proxy variable for mental health, allowing researchers to estimate the prevalence of mental health problems in children with specific medical conditions or from specific populations.

In terms of medical conditions, over the past four years the PSC has been used to estimate the prevalence of mental health problems in three studies in children with asthma (Lee et al., 2019; McGovern et al., 2019; Pesek et al., 2021), two in burn injuries (Riobueno-Naylor et al., 2019; Riobueno-Naylor et al., 2021), two in cancer patients or survivors (Groves et al., 2019; Michaud et al., 2020), one in celiac disease (Chellan et al., 2019), one in diabetes (Semenkovich et al., 2019), two in epilepsy (Afzal et al., 2021; Choudhary et al., 2017), two in functional constipation (El-Sonbaty et al., 2019; Klages et al., 2017) , one in hematopoietic cell transplantation (Buchbinder et al., 2019), one in hip preservation surgery (Richard et al., 2020), two in HIV (Joshi et al., 2017; Kefale et al., 2019), one in menarche (Irannezhad & Soltanizadeh, 2021), one in pediatric hospitalizations (Doupnik et al., 2017), one in restless legs syndrome (Stubbs & Walters, 2020), three in sleep difficulties (Basch et al., 2019; Kim et al., 2017; Lin et al., 2021), three in thalassemia (Elzaree et al., 2018; Ghorbanpoor et al., 2020; Raman et al., 2019), two in substance use disorders (Anyanwu et al., 2017; Shahzad et al., 2020), four in weight management (Gowey et al., 2020; Ratcliff et al., 2018; Thomaseo Burton et al., 2020; Wylie-Rosett et al., 2018), and three in chronic conditions in general (Hastuti et al., 2020; Im & Kim, 2021; Rukabyarwema et al., 2019). An additional nine studies have used the PSC to explore the effects of the COVID-19 pandemic on psychosocial functioning (Bate et al., 2021; Bourne et al., 2021; Hussong et al., 2021; Lund & Gabrielli, 2021; Mekori-Domachevsky et al., 2021; Oliva et al., 2021; Ozturk & Ayaz-Alkaya, 2021; Riffat et al., 2021; Shah et al., 2021), and fourteen studies have used the PSC to learn more about how hunger, poverty, victimization, or other adverse childhood experiences affect psychosocial functioning (Burdzovic Andreas & O'Farrell, 2017; Choi et al., 2020; Fonseca et al., 2019; Gaete, 2021; Hawkins, 2020; Leiva et al., 2021; López et al., 2018; Metwally et al., 2020; Nasir et al., 2019; Roby et al., 2021; Suku et al., 2019; Vargas et al., 2019; Weigel & Armijos, 2018).

All three of the PSC subscale scores have been used as proxies or screens for diagnosis-related mental health conditions in pediatric primary care settings. In addition to the already mentioned studies of the PSC Externalizing scale by Kolko and his associates (Kolko et al., 2014; Kolko et al., 2010; Kolko et al., 2011; Kolko et al., 2020), Jellinek and colleagues (2021) found that using the PSC-17P internalizing scale (whether or not in conjunction with the PSC overall scale score) as a first-stage screener identified most adolescents who were screened at-risk on the PHQ-9M, supporting the PSC's use as a depression screen as well as an overall screen. Spencer and colleagues (2018) also found that the PSC attention subscale could accurately identify ADHD in a clinical population and that the overall and other subscales could accurately distinguish between complex and simple presentations of ADHD. The PSC has also been used to validate new screening measures, including a trauma screen (Sachser et al., 2017) and an early childhood psychosocial screening tool (Fallucco et al., 2017).

In terms of special populations, Tolliver and colleagues (2020) used the PSC to assess the prevalence of mental health problems in rural primary care, while Spencer and colleagues (2019) used the PSC as a proxy for overall mental health problems to study referrals to embedded child psychiatry in an urban, Latino primary care population. In an urban primary care setting, Mehus and colleagues (2019) found that parents whose children had PSC risk expressed greater interest in accessing resources to address their children’s mental health. Using the PSC, other studies have found that children’s mental health is associated with parental mental health (Huang et al., 2018), including in military/veteran families (Zalta et al., 2018). In other studies, heightened prevalence of PSC symptoms in a particular community has validated the need for interventions for that group, such as in the case of a parenting intervention in North Carolina (Baugh et al., 2019) and a school-based trauma intervention for immigrant youth (Mancini, 2020).

Although the following studies have already been cited with regard to their main focus, they can also be grouped together as studies of screening in languages other than English in other countries or in immigrant populations in the US. This global use of the PSC can be viewed as another indicator of feasibility. The PSC has been used in non US settings such as Chile (Gaete, 2021; Leiva et al., 2021; Leiva et al., 2019; López et al., 2018; Vargas et al., 2019), Ecuador (Weigel & Armijos, 2018), Egypt (El-Sonbaty et al., 2019; Elzaree et al., 2018; Metwally et al., 2020), Ethiopia (Kefale et al., 2019), Ghana (Huang et al., 2018), India (Chaurasiya et al., 2019; Chellan et al., 2019; Choudhary et al., 2017; Gupta et al., 2017; Raman et al., 2019; Shah et al., 2021), Indonesia (Riffat et al., 2021), Iran (Ghorbanpoor et al., 2020; Irannezhad & Soltanizadeh, 2021), Nigeria (Anyanwu et al., 2017), Pakistan (Shahzad et al., 2020), Rwanda (Rukabyarwema et al., 2019), Japan (Higuchi et al., 2021) and Turkey (Ardıç et al., 2020; Öner et al., 2019; Ozturk & Ayaz-Alkaya, 2021).

Literature reviews mentioning the PSC in papers related to pediatric screening (see Appendix 1, Section II)

There were 11 reviews comparing measures that could be used for psychosocial screening in pediatric settings, of which 10 ultimately recommended the PSC. Three of these studies exclusively recommended the PSC. Pourat and colleagues (2017) and Zima and colleagues (2019) both reported on a UCLA review of 15 child mental health screeners in order to select a clinical outcome measure for California's publicly funded outpatient mental health programs. Their review used provider surveys, scientific literature reviews, a modified Delphi panel, and ratings based on nine criteria, including effective care, scientific acceptability, usability, feasibility, and overall utility. The PSC-35 "rose to the top based on acceptable scientific evidence and high Delphi panel ratings and was the only measure that met all nine minimum criteria" (Zima et al., 2019, p. 385). In another review of 16 screening measures to identify adolescent risk-taking behavior, Hiott and colleagues (2018) reported that "When viewed through the Donabedian Framework of structure process and outcome, the Pediatric Symptom Checklist and its derivatives are the screens of choice."

Seven additional studies recommended both the PSC and other measures. Marlow and associates (2019) reviewed tools for screening for mental health problems in children with developmental disorders in low- and middle-income countries, and endorsed the PSC because it is freely available, has specificity and sensitivity above 70%, and has been used in non-Western settings. Out of 10 child clinical outcome measures reviewed by Marti and colleagues (2021), the PSC was one of the three measures recommended because its psychometric properties have been tested across diverse populations. Caballero and colleagues (2017) recommend that primary care providers consider using the PSC particularly to meet the mental health needs of Latino children in immigrant families, noting that using the pictorial version of the PSC can be especially advantageous in low-resource, low literacy settings, and that the PSC is also valid in Spanish for these families. Several other reviews supporting the use of the PSC have emphasized its validity as a broadband screening tool, noting its "excellent" classification accuracy (Cianchetti, 2020) and general psychometric support (Becker-Haimes et al., 2020; Jeffrey et al., 2020; Trafalis et al., 2021).

The PSC has been referenced or recommended for clinical use by other review papers and commentaries because it is a well-known, common, and valid measure for overall psychosocial screening purposes (Acra & Perez, 2017; Black & Rofey, 2018; Brino, 2020; Burt et al., 2018; D'Angelo et al., 2018; Forman-Hoffman & Viswanathan, 2018; Joshi & Apple, 2019; Simha & Brown, 2021). Based on all these studies, it seems clear that the evidence supporting the use of the PSC as a psychosocial screen for routine use in pediatrics is now even stronger than it was four years ago when the evidence was last reviewed by NQF.

Stepping back from the many supporting studies, we think it is important to acknowledge that there is still a lack of large scale, population-based evidence for the impact of universal screening for overall psychosocial problems in pediatric populations. This is probably the main reason that the US Preventive Services Task Force has yet to endorse it. It is, however, important to note that this lack is not unique to this field. Quite the contrary, recent papers commissioned by NIMH and NICHD (Gardner et al., 2021; King et al., 2021) have pointed out the fact that the current limited evidence for universal screening in pediatrics applies to most medical—as well as to most mental health—problems. And yet, the same papers reflect a continuing commitment to screening in general and to screening for psychosocial problems in particular.

Evidence for this continuing commitment is a special issue of *Pediatrics* in July 2021 that was devoted solely to this topic. The lead paper in this issue is by William Gardner and his associates and its title is probably the best simple summary of the entire field. Using the USPSTF recommendations for universal screening for adolescent depression as a case in point, the paper points out the "Potential for Population Health Effectiveness of Screening", clearly showing the continued relevance as well as the need for improvement in such screenings.

We wanted to conclude our proposal for NQF re-endorsement of psychosocial screening with the PSC tool by attempting to put the issues in a broader context. As the *Pediatrics* special issue on screening makes clear, when it comes to universal screening in pediatrics, there is much to learn. Since the diagnostic system for children's emotional disorders is at best too broad and non-specific, as well as too fragmented into hundreds of supposedly discrete but actually overlapping diagnoses, most of the work to date on mental health screening has recommended that the focus should be on assessing functioning in key areas of daily life (family, school, friends, activities and mood) instead of screening based on formal diagnostic categories.

Although we do not know the long-term benefits of screening, we do know clinically that helping children with emotional needs probably enhances or eases ongoing deterioration of their self-esteem and reduces their suffering. We also know that many if not most parents are eager for advice from their pediatricians and/or mental health experts concerning their child's emotional functioning and developmental issues. At this point, the PSC is the instrument most widely used in the real world of clinical settings. It focuses on psychosocial functioning, can interface with diagnostic criteria, recognizes emotional issues in an efficient manner, can quantify the severity of problems found, and provides a platform for discussion between pediatrician and parent. These qualities, along with the PSC's statistical accuracy and widespread use in pediatrics, argue strongly for its continued endorsement by NQF.

1a.15) Detail the process used to identify the evidence.

Literature review

1a.16) Provide the citation(s) for the evidence.

References

A list of all PSC related studies published since our 2017 NQF proposal is provided in Appendix 1 with studies grouped in six categories according to their primary focus. A full list of the references cited in our 2017 proposal, along with the PSC papers published since 2017 and all of the other papers cited in the current proposal is available in Appendix 2.

[Response Ends]

1a.15. Detail the process used to identify the evidence.

[Response Begins]
2021 Submission
Literature review and expert option provided by AAP Task Force on Mental Health
[Response Ends]

1a.16. Provide the citation(s) for the evidence.

[Response Begins]
2021 Submission
American Academy of Pediatrics. The Case for Routine Mental Health Screening. *Pediatrics.* 2010, **125,** s133-139.

American Academy of Pediatrics. AAP Publications Reaffirmed or Retired. Pediatrics. 2021;147(6):e2021051488. https://doi.org/10.1542/peds.2021-051488.

[Response Ends]

Performance Gap

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]
Psychosocial problems in children are common and treatable with prevalence estimates of about 12% of all children and adolescents (Gardner, Lucas, Kolko, & Campo, 2007; Kelleher et al., 1997; Murphy et al., 2016). Studies have shown that children with these problems are often unrecognized by their pediatricians (~50% of cases) (Kelleher et al., 1997) and that only a fraction of them receive treatment (Hacker et al., 2014b; Kelleher et al., 1997). Children with psychosocial problems are more likely to have poorer health, academic, behavioral, and social outcomes in both the short and long term (Murphy et al., 2015). Children who receive psychosocial screening as a part of pediatric well child visits are more likely to receive outpatient mental health services (Hacker et al., 2014a; Hacker et al., 2014b; Savageau et al., 2016) than are children who are not screened. As the dates of the studies just cited attest, it is only within the last three years that strong evidence documenting the relationship between psychosocial screening and increased mental health treatment has become available.

A series of RCT studies by Kolko and his associates have shown that pediatric outpatients with a wide range of problems who are found to be at risk when screened with the PSC and go on to receive pediatric office based mental health interventions have significantly lower mental health symptom scores and better functioning at immediate and longer term follow up than do similar outpatients randomized to treatment as usual (Kolko et al., 2014; Kolko, Campo, Kelleher, & Cheng, 2010). For these reasons, we believe that an increase in mental health treatment is the most appropriate (and a measurable) benchmark for assessing the positive impact of routine psychosocial screening. The logic model for screening in pediatrics is that more children will receive help, fewer children will develop mental, emotional, and behavioral disorders (Guzmán et al., 2015; Kieling et al., 2011), and more children who received help will enjoy better life outcomes (Kellam et al., 2014).

Requiring screening for psychosocial problems as part of routine well child care in pediatrics is one of the most frequently recommended ways to improve recognition and intervention for such problems (Hacker et al., 2014a) and an increasing number of states (Massachusetts (Savageau et al., 2016)), insurers (Medicaid/EPSDT (Mann, 2013)), standard setting organizations (American Academy of Pediatrics (Foy, Kelleher, Laraque, & Health, 2010; Weitzman & Wegner, 2015)), blue ribbon panels (President's New Freedom Commission on Mental Health (Hogan, 2003) (Institute of Medicine (O´Connell, Boat, & Warner, 2009)), and advocacy organizations such as the Kennedy Forum (Fortney et al., 2015) and Mental Health America (http://www.mentalhealthamerica.net/positions/early-identification) have now required, endorsed, or recommended the principle of including a psychosocial screen as a part of every well child visit for children aged 3-17.

The PSC is probably the most frequently recommended and widely used tool for routine psychosocial screening in pediatrics (Semansky, Koyanagi, & Vandivort-Warren, 2003), with the Strengths and Difficulties Questionnaire (Goodman, Meltzer, & Bailey, 1998) and Child Behavior Checklist (Achenbach, 2009) instruments that are similar in many ways and also frequently mentioned and validated in this context. Many of the endorsements noted above include these three and/or a few other instruments.

The reference list is included in the attached appendix.
[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]
1b.2. Provide performance scores on the measure as specified (*current and over time*) at the specified level of analysis. (*This is required for maintenance of endorsement*. Include mean, std dev, min, max, interquartile range,

scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include). This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

2017 Submission

Several studies based on data from the Massachusetts Medicaid (MassHealth) pediatric behavioral health screening program and the Children's Behavioral Health Initiative (CBHI), demonstrate the currently wide variation in the rates of mental health screening with formal, standardized tools as well as the possibility of improvement and the potential benefits of doing so.

2021 Submission

In this section, we begin with the material we presented in our 2017 NQF request for endorsement (unchanged, in black font) with new material presented in red font.

Data source 1:

Summary data from the CBHI Behavioral Health Screening Cumulative Quarterly Report; posted on BHSCQR website

Measurement Period dates of service from 1/1/2008 to 3/31/2017

Data Source 1a: Statewide data for all children

Denominator/Well child visits for all children .5 -20 years of age: 4,721,790

Numerator (screens with visit): 2,965,923

Statewide average: 62.8%

Minimum: 14.2%

Maximum: 71.9%

Standard Deviation: 12.4%

95% Confidence interval: 58.8% to 66.8%

The CBHI BHSCQR also presents the same statewide Medicaid screening data broken down for each of the state's regions for each of the 37 calendar quarters from Q1 (January through March of 2008) to Q37 (January through March of 2017) of CBHI. Table 1 (below) reports the rates of screening in the state's six regions with data from the four quarters in each year averaged for simpler presentation in this proposal. Because the BHSCQR website does not break out the data by age group, the information in the table below is for the full sample of all ages, but the trends should be large enough that that the patterns shown below should be quite similar to those for just the 3–17-year-olds. As the table shows, the range in rates of screening vary widely, from a low of 21.55 in the Boston region in 2008 to a high of 85.27 in the Western region in first quarter of 2017. The percentages of well child visits with screens goes up substantially in all regions over the first few years of the initiative but after that the rank ordering remains relatively consistent across regions. Western Massachusetts always has the highest rate of screening, Metro West and Boston the lowest and Northeast, Southeast, and Central Massachusetts in the middle.

Table 1. Rates of Screening in different regions of Massachusetts from 2013-2017 by Region

Region	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017***
Western	38.14%	66.13%	75.34%	79.16%	79.79%	80.33%	81.62%	83.76%	81.27%	85.27%
Central	32.61%	51.99%	60.61%	63.22%	67.23%	70.59%	71.09%	70.90%	72.33%	72.31%
Northeast	32.70%	56.36%	60.93%	60.31%	55.35%	65.08%	67.16%	68.43%	71.46%	67.38%
Metro West	25.89%	47.45%	51.92%	51.80%	57.72%	60.47%	59.83%	54.90%	49.84%	47.72%
Southeast	36.88%	61.89%	70.57%	75.94%	78.18%	80.19%	80.66%	77.87%	75.91%	73.95%
Boston	21.55%	45.91%	55.28%	57.26%	60.53%	64.49%	67.57%	65.12%	63.27%	58.76%

Note. Rates of screening were calculated by the year from 2008-2016 and for the first quarter of 2017

*Data available only for the 1st quarter of 2017

Table 2 below presents the distribution of rates of screening in greater detail, taken from this time directly from BHSCQR website for all 222 measurement points (37 quarters in all six regions), illustrating even more dramatically the wide range of rates of screening across the state and over nearly a decade.

Table 2 The distribution of screening rates by decile broken down by quarters for each region from January 2008 through March 2017 (6 regions x 4 quarters x 9.25 years) for children of all ages
<u>2021 Submission</u>

% of Well-Child Visits with Screens (in Deciles)	**# of Quarters in Decile**
0-9%	1
10-19%	5
20-29%	4
30-39%	8
40-49%	23
50-59%	40
60-69%	63
70-79%	58
80-89%	20
90-100%	0
*	*
Total (Regions*Quarters*Years)	222
*	*
Min:	8.21%
Max:	85.65%
Range:	77.44%
Median:	64.53%

*Cell intentionally left blank

***Data Source 1b*:** Statewide data broken down by for children ages 3-17 years

Denominator /Well child visits for all children 3 -17 years of age: 2,361,475

Numerator (screens with visits): 1,681,764

Statewide average: 71.2%

Minimum: 39.98%

Maximum: 79.14%

Standard Deviation: 11.3%

95% Confidence Interval: 64.2% to 78.2%

We included data source 1b focusing just on the 3-17 year old children in the BHSCQR website data as this is the age group that was screened with the PSC, and prior studies (Hacker et al., 2016; Savageau et al., 2016) have shown that the PSC was the measure used for 67% of the children in the CBHI 3-17 year old age group. By multiplying the total number of screens in this age group by 67%, we can estimate that approximately 1,126,782 PSC's were administered over the first 9.25 years of CBHI.

Table 3 (below) shows the Massachusetts statewide number of well child visits, number of screens, and percent of visits with screens, for just the 3–17-year-old (PSC screened) children with Medicaid from January 2008 (start of CBHI) to March of 2017. Data in this table are taken directly from the CBHI BHSCQR but with the four quarters of each year aggregated together so that the totals for each year could be seen more clearly. As the table shows, the rate of screening rose from approximately 39.98% for its first year (2008) to 65.72% for its second year to over 70% for its third year, and then remaining in the 70% range in all of the six years since. Not shown in this table but present in the data shown on the CBHI BHSCQR are the figures for the first quarter of 2008 (which show a base rate of 17.8% during the first three months of the program). Although not posted on the CBHI BHSCQR but reported by two different groups (Hacker et al., 2016; Savageau et al., 2016) with access to claims data for 2007 is that the rate of formal screening during the year prior to the start of CBHI was less than 5%. A rate of formal screening that started and then remained at less than 5% was also reported for the state of California (which had no requirement for the use of formal screens) for 2008 and 2009 when the rate in Massachusetts had climbed to about 65%.

Table 3. Statewide rates of formal psychosocial screening for 3–17-year-olds during WCV from 1/1/2007 to 3/31/2017

Year	**Denominator** (Total Visits)	**Numerator** (Total with Screens)	**% Visits with Screens**
12/31/07-12/31/2008	202,376	80,910	39.98%
1/1/2009-12/31/2009	219548	144276	65.72%
1/1/2010-12/31/2010	234823	171114	72.87%
1/1/2011-12/31/2011	243209	179305	73.72%
1/1/2012-12/31/2012	252357	190286	75.40%
1/1/2013-12/31/2013	265826	208549	78.45%
1/1/2014-12/31/2014	287690	228467	79.14%
1/1/2015-12/31/2015	290313	217282	74.84%
1/1/2016-12/31/2016	300307	215611	71.80%
1/1/2017-3/31/2017	65,026	45,964	70.69%
Total	2,361,475	1,681,764	71.22%

*Cell intentionally left blank

2021 Submission
Data Source 1x: Site level data from CBHI for pediatric patients in NHP (a Medicaid HMO)

Measurement Period dates of service from 1/1/2016 to 12/31/2016

Denominator /Well child visits for all children aged 4.0 to 17.99: 76,021

Numerator (screen at WCV): 57,321

Average: 75.4%

Min: 0%

Max: 100%

Standard Deviation: 12.7%

This data source was actual claims data for a subsample of Data Source 1 described above, in this case of the WCV of all 4-17 year old children from practices that submitted claims for at least 30 cases who were covered by NHP, the largest Medicaid HMO in Massachusetts.

2017 Submission
Data Source 2: CBHI Cohort Data for chart review sample

Measurement Period dates of service in 2007, 2010, and 2012

Denominator /Well child visits for all children: 4,977

Numerator (screen at WCV): ~1700

Average: 51.7%

Min: 1.5%

Max: 88.9%

Standard Deviation: 35.1%

95% Confidence interval: 35.5% to 87.2%

This second data source is a ~ 6000 visits subsample of the CBHI statewide data retrieved from chart reviews supplemented by administrative claims data.

Differences in rates of screening in three cohorts of about 2000 cases of pediatric outpatients (age 4-17 years) seen before (2007) and after (2010, 2012) the start of CBHI. As shown in Table 4, consistent with the statewide CBHI data, the Southeast and Western regions consistently demonstrated significantly higher rates of screening, the Northeast and Metro West demonstrated the lowest rates of screening, and Central Massachusetts and Boston demonstrated screening rates in the middle.

Table 4. Rates of Formal Screening for each year for children ages 4-17

-

	Formal Screen in 2007		**Formal Screen in 2010**		**Formal Screen in 2012**	
	No	**Yes**	**No**	**Yes**	**No**	**Yes**
Region		*[1]		***[2]		***[3]
Western	139 (95.2%)	7 (4.8%)	22 (11.9%)	163 (88.1%)	33 (15.2%)	184 (84.8%)
Central	64 (94.1%)	4 (5.9%)	30 (23.3%)	99 (76.7%)	31 (23.9%)	99 (76.1%)
Northeast	125 (91.2%)	12 (8.8%)	63 (37.5%)	105 (62.5%)	71 (39.0%)	111 (61.0%)
Metro West	84 (94.4%)	5 (5.6%)	54 (32.5%)	112 (67.5%)	62 (37.4%)	104 (62.7%)
Southeastern	197 (98.5%)	3 (1.5%)	52 (23.0%)	174 (77.0%)	27 (11.1%)	217 (88.9%)
Boston	149 (98.0%)	3 (2.0%)	51 (25.4%)	150 (74.6%)	33 (18.2%)	148 (81.8%)

*Cell intentionally left blank

2007: $^{1}x^{2}=13.33$ $p=.021$ **2010**: $^{2}x^{2}=36.33$ $p<.001$ **2012:** $^{\mathbf{3}}x^{2}=75.21$ $p<.001$

Data Source 3: Medicaid screening data from four Massachusetts General Hospital outpatient clinics

Measurement Period dates of service from 7/1/2014 to 12/31/2016

Denominator /Well child visits for all children 4-17 years of age: 10,334

Numerator (screen at WCV): 7,915

Average: 76.6%

Min: 9.4%

Max: 91.7%

Standard Deviation: 38.5%

95% Confidence interval: 38.8% to 100.0%

One of the most important criteria for a quality measure is differences in performance across sites. Although the data from the CBHI BHSCQR summarized above provide strong evidence that there are consistent differences in screening rates for different regions of the state and different age groups, our lack of access to the actual data made it impossible to explore differences between different pediatric practices, one of the intended uses of this type of quality measurement. Since it was not possible for us to obtain the actual data from the state of Massachusetts and since the sample from Data Source 2 was too small to permit analyses by clinic, we turned to our own hospital system and were able to obtain billing data for pediatric psychosocial screens for a relatively large sample of well child visits in four MGH-affiliated outpatient pediatric clinics. Using the same CPT billing code for screening (96110) used by the state and focusing only on children with Medicaid health insurance, we obtained data on 10,827 children aged 4-17 who had at least one pediatric well child visit from July 2014 through December 2016 in one of the four clinics.

Differences in rates by clinic: As shown in Table 5 below, the differences in rates of mental health screening at well-child visits were substantial and statistically significant between these four clinics. In 2016, for example, rates ranged from 0% to 88.1% (p < .001) to in the four clinics. Over the three years of data, the rank orders of screening rates among the four clinics were relatively constant with Clinic A and B as always the highest and C the lowest in each of the three years (not shown).

Table 5: Rates of screening in four MGH affiliated outpatient pediatric clinics in 2016

Variable	Level	Overall*	2016 (n=6,801 children): January 1, 2016 – December 31 2016				
			Clinic A	**Clinic B**	**Clinic C**	**Clinic D**	**p-value over 4 sites**
Billed for MH screen	**Yes**	69.7% (4743)	86.1% (3422)	88.1% (1215)	0.0% (0)	24.1% (106)	<.001

2021 Submission

***Data Source 4*: Data from a predominantly commercially insured sample of WCV with electronically administered PSC's**

Measurement Period dates of service from 5/1/2018 to 12/31/2019 (excluding pilot phase)

Denominator /Well child visits for all 5.5-17.99 year old children: 45,949

Numerator (screen at WCV): 39,202

Average: 85.3%

Min: 70.2%

Max: 97.4%

Standard Deviation: 12.7%

This data source was electronically administered PSC form data from a network of 15 practices for all 5.5-17.99 year old children seen for WCV in 2018 and 2019.

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.

[Response Begins]
N/A
[Response Ends]

1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]
Differences by age:
As presented below in Table 6 from Data Source 1, the CBHI BHSCQR shows significant differences in rates of screening by age group with very young (< 3) children and older (> 17) youth less likely to be screened than 3–17-year-olds. Even within the 3–17-year-old group there were differences in screening rates with younger (3–6-year-olds) and older (13- to 17-year-olds) patients showing lower rates of screening than 7–12-year-olds (72.41% vs 67.32% vs 73.08%) respectively.

Table 6 (December 31, 2007 - March 31, 2017)

Age	Total Visits	Total Visits with Screens	%Visits with Screens
<6 mos to 2 years	**2,136,135**	**1,205,607**	**56.44%**
<6 mos	882,434	336,179	38.10%
6 mos to 2yrs	1,253,701	869,428	69.35%
3 yrs to 17 yrs	**2,361,475**	**1,681,764**	**71.22%**
3 yrs to 6 yrs	776,570	562,311	72.41%
7 yrs to 12 yrs	911,693	666,266	73.08%
13 yrs to 17 yrs	673,212	453,187	67.32%
18 yrs to 20 yrs	**224,180**	**78,552**	**35.04%**
Total	4,721,790	2,965,923	62.80%

Lack of disparities by race, ethnicity, and language:

Although none of the published papers on the CBHI sample (Hacker et al., 2014a; Hacker et al., 2014b; Savageau et al., 2016) explore data on disparities by race, ethnicity, or language, the chart review study of a subsample of these cases from Data Source 1 (Savageau et al., 2016; Savageau, Simons, Lucke, Jellinek, & Murphy, 2017) explored screening by demographics in a subsample of ~ 6000 visits from 2007, 2007, and 2010. As shown in Table 7 below, there were no significant disparities by race, ethnicity, or language. It may be important to note that disparities by socioeconomic status cannot be meaningfully assessed in this sample since by CBHI is a program only for children with Medicaid and SES is confounded with insurance type.

2021 Submission
Table 7. Lack of Significant Disparities of Children 4-17 years insured by MassHealth, FYs 2007*, 2010, and 2012

		Had a Formal Screen in 2007	Had a Formal Screen in 2010	Had a Formal Screen in 2012
Race	**White**	7 (2.9%)	374 (73.6%)	397 (73.7%)
	Non-White	20 (6.4%)	276 (77.7%)	290 (79.2%)
Ethnicity	**Non-Hispanic**	17 (4.4%)	321 (75.3%)	369 (78.7%)
	Hispanic	10 (5.8%)	224 (79.4%)	246 (80.9%)
Primary Language	**English**	24 (3.8%)	676 (75.4%)	716 (76.8%)
	Non-English	10 (6.5%)	128 (71.1%)	149 (76.4%)

*Cell intentionally left blank

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.

[Response Begins]
N/A
[Response Ends]

2. Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.

---

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see What Good Looks Like).

[Response Begins]
Psychosocial Screening Using the Pediatric Symptom Checklist-Tool (PSC-Tool)
[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]
Percentage of children from 3.00 to 17.99 years of age seen for a pediatric well child visit who have a Pediatric Symptom Checklist (PSC) Tool administered as a component of that visit.
[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]
Behavioral Health
[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]
Health and Functional Status
Screening
[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]
Children (Age < 18)
[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]
Facility
Health Plan
Population: Regional and State
[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.
[Response Begins]
Outpatient Services
[Response Ends]

sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.

Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".

[Response Begins]
https://www.massgeneral.org/pediatric-symptom-checklist/
[Response Ends]

sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.
[Response Begins]
No data dictionary/code table – all information provided in the submission form
[Response Ends]

sp.12. State the numerator.

Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).

DO NOT include the rationale for the measure.

[Response Begins]
Number of patients with documentation that the PSC tool was administered as part of the well child visit.
[Response Ends]

sp.13. Provide details needed to calculate the numerator.

All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]
Patients passing this quality measure are identified through a review of the medical record. In a chart review, the presence of a PSC score or PDF scan of it in the progress note, or score shown in the visit template or flowsheet documents the completion of the screen on the same day of the WCV. To receive credit, progress notes must indicate the name of the specific measure and actual score (e.g., PSC given, score = not at risk).

2021 Submission
In pediatric settings in which the PSC is completed electronically (e.g., over the internet or on an iPad in the waiting room), item- and subscale scores are filed and available in defined fields in the EHR (such as flowsheet or visit note template) and the numerator is established by checking these fields for data on the date of the WCV.

There are four versions of the PSC in wide clinical use: the original 35 item form that is filled out by parents, a 17 item version of this, and 35 and 17 item versions of a youth report version. All four of these versions have the same 17 items at their core and the same subscale scores. Credit for administering the screen is given if any of the four have been used.

In Massachusetts (and possibly in other locations which require routine screening), providers who conduct well child visits with patients covered by Medicaid are required to bill for each screen using CPT code 96110 or 96127 so that officials can ascertain whether a screen was given. Therefore, in the Massachusetts program, the presence of CPT code 96110 or 96127 (in the billing data or in the EHR) can be used to identify patients in the numerator with the denominator being all cases with a CPT code for a WCV for a given time period (e.g., calendar year) and entity (e.g., specific provider, health center, health plan, etc.).
[Response Ends]

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]
Number of patients aged 3.00 to 17.99 seen for a pediatric well-child visit.

[Response Ends]

sp.15. Provide details needed to calculate the denominator.

All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]
Cases are identified from administrative data for site. Number of unique patients ages 3.00 to 17.99 seen for a well-child visit (CPT 99381-99394) in a defined evaluation period, often a year.
[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]
No exclusions.
[Response Ends]

sp.17. Provide details needed to calculate the denominator exclusions.

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]
N/A
[Response Ends]

sp.18. Provide all information required to stratify the measure results, if necessary.

Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.

[Response Begins]
N/A
[Response Ends]

sp.19. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.
[Response Begins]
No risk adjustment or risk stratification

[Response Ends]

sp.20. Select the most relevant type of score.

Attachment: If available, please provide a sample report.
[Response Begins]
Rate/proportion
[Response Ends]

sp.21. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score
[Response Begins]
Better quality = Higher score
[Response Ends]

sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]
Step 1. Count number of children aged 3-17 seen for a well child visit in state, region, clinic, or other group during defined period (often, one year) using administrative data (CPT 99381-99394). N=total population. This is the denominator.
Step 2. Assess whether PSC was administered as a part of WCV, for the eligible population, using the chart for indicator status. Pass if documentation that screen was given on the day of the WCV is present.
Step 3. Compute numerator = count of patients with completed PSC.
Step 4. Calculate clinic or other entity rate as numerator/denominator. No risk adjustment.
[Response Ends]

sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

[Response Begins]
This measure and its denominator are not based on samples. This PRO measure is based on a parent or child completing the PSC (no proxy) and noting its presence/absence. Missing data (no administration of PSC) is managed by the inclusion of patients without a completed PSC in the denominator.
[Response Ends]

sp.28. Select only the data sources for which the measure is specified.

[Response Begins]
Claims
Electronic Health Records
Paper Medical Records
[Response Ends]

sp.29. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]
2017 Submission
In medical record (paper or electronic):
If patient age => 3.0 & age =< 17.99; claim for well child visit (99382 or 99383 or 99385 or 99392 or 99393 or 99394), assess progress note, templated note, flowsheet, scanned in PSC, for evidence that screen was administered.

2021 Submission
In systems like the state of Massachusetts where billing for the screen as well as screening itself is required, the source for data about screening can be claims data.
In both our 2021 and 2017 submissions we stated that paper medical records were a possible source of measure information for which evidence was available. Data source 2 was based on a review of ~4000 (in some comparisons ~ 6000) paper medical records. Some information about the use of this sample to assess ethnicity and regional differences in the statewide CBHI program is provided above in section 2a.08. As described in the Savageau, et al 2016 paper, the reliability and validity of this method was established in the following way: "An experienced medical record review (MRR) vendor conducted a retrospective MRR of the sample's ~4000 children/adolescents. Three registered nurses used a chart abstraction tool developed by one of the researchers and a panel of practicing physicians. Before implementation of the MRR, nurse abstractors had to pass Gold Standard Testing and attain interrater reliability scores of 95% or higher. The abstraction tool was piloted in a large community-based practice. From chart notes and documentation, nurse abstractors determined the presence of standardized BH screening, screening results, and referrals. They also detailed the presence of non-MassHealth approved screening tools and notations of informal screening/surveillance without a specific tool. Where both formal and informal screens were conducted and abstracted, subsequent analyses prioritized results from the formal screening. In addition, abstractors recorded charted notes and documentation on BH referrals (made at the time of the WCV), patient demographics (*i.e.*, age, sex, ethnicity, and primary language spoken at home), interpreter use during the WCV, and use of a non-English BH screening tool.

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]
Available at measure-specific web page URL identified in sp.09
[Response Ends]

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

[Response Begins]
Yes
[Response Ends]

2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing.

Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

[Response Begins]
Yes
[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]
No
[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]
No additional risk adjustment analysis included
[Response Ends]

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.
2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure

score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.
2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument-based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.
2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;
AND
If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).
2b3. For outcome measures and other measures when indicated (e.g., resource use):
• an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration
OR
• rationale/data support no risk adjustment/ stratification.
2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;
OR
there is evidence of overall less-than-optimal performance.
2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.
2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.
2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:
2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and
2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.
(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions
Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include but are not limited to inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.
Examples of evidence that an exclusion distorts measure results include, but are not limited to frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
Risk factors that influence outcomes should not be specified as exclusions.
With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v.$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Scientific Acceptability sections. For example:

2021 Submission:
Updated testing information here.

2018 Submission:
Testing from the previous submission here.

Reliability

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]
Claims
Electronic Health Data
Electronic Health Records
[Response Ends]

2a.02. If an existing dataset was used, identify the specific dataset.

The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).

[Response Begins]
2021 Submission
Psychosocial screening continues to be a required component of well-child visits (WCVs) for all children and adolescents with Medicaid in a statewide program in Massachusetts called the Children’s Behavioral Health Initiative (CBHI). The PSC-Tool is the measure used to screen the majority (67%) of children aged 5.5 to 17.99 years of age in that state (Savageau, et al, 2015). As of the end of September 2020 (the last quarter posted), when this program had almost completed its 13th year, more than 4.5 million screens had been administered, with just over half of them (2.5 M) with 3–17-year-olds and therefore an estimated 1.7 million of them being PSC's. It may also be important to note that although the website shows that the number of WCV and screens dropped considerably during the second quarter of 2020 due to the COVID-19 pandemic, by the next quarter (the most recent available) both rates had rebounded to near their pre-pandemic levels. Anecdotal data from clinicians suggests that by the end of 2020 and in the first half of 2021, rates of WCV and screening had remained at or near their previous levels.

In any case, we had based our 2017 NQF submission on three subsamples from the CBHI data set with data from 2008 to the end of the first quarter of 2017 and did not update it with the more recent data in our current NQF submission. Instead, for this new, (Fall 2021) submission, we were able to obtain actual claims data for a subsample of the cases in our original **Data Source #1**, so we have added a description of this subsample as **Data Source 1x** in the sections below. This data set allowed us to explore differences at the site level in rates of psychosocial screening using claims data.

Psychosocial screening with the PSC-Tool is also a required component of WCV's in a number of different networks of outpatient pediatric practices in Massachusetts, providing researchers with first time ever data on the results of screening exclusively with the PSC, as well as in practices that serve primarily commercially insured children. Since we were able to gain full access to all of the screening data over two years from one of these networks, we have added it as **Data Source 4.** This data set allowed us to assess differences across sites in a 15-site network and to test reliability and validity of data on PSC administration, that was collected, scored, and stored in defined fields in patients' EHR's.

The descriptions of **Data Sources 1, 2, and 3**, appear in black font below, unchanged from what we wrote in our 2017 NQF submission. Descriptions of **Data Sources 1x and 4** are new with this submission and appear in red font.

2017 Submission
Data Source 1:

Summary Data from CBHI pediatric behavioral health screening website

MassHealth posts a running summary of administrative claims data (Behavioral Health Screening Cumulative Quarterly Report; BHSCQR) for all well child visits for children with Medicaid in Massachusetts since the start of CBHI in January 2008 on a publicly accessible website (https://www.mass.gov/info-details/cbhi-data-reports#behavioral-health-(bh)-screening-cumulative-quarterly-report-). The CBHI BHSCQR summarizes the number of WCV and the number and percentage of these visits with behavioral health screens and several other variables over the course of the initiative. Data are presented for the state as a whole and also broken down by age group and region of the state.

2021 Submission
For the calculations in our 2017 submission, we were limited to the state's published data going up to the end of the first quarter of 2017. Although the state website now shows data until the end of 2020, we have not updated the data we presented in 2017, except as noted below.

Data Sources 1, 1x, 2, and 3 are all from the state of Massachusetts CBHI program and all of the data sources rely on the same administrative claims data. In our 2017 NQF submission Measure Testing Form, we reported on the reliability and validity of this data in Data Source 3 only but as representing all CBHI data sources, comparing results based on claims data to chart review and inter-coder reliability as the gold standards. The material that we presented in the MTF in 2017 can be found in the descriptions of Data Source 3 below in black font.

Data Source 1x:

Over the three years since our previous NQF submission, we were able to gain access to the actual administrative claims data and demographic information for a subset of the data reported in Data Source 1 (not just the summary data that we had previously reported on). Neighborhood Health Plan (now called Allways Health) is the largest Medicaid HMO in Massachusetts. We received de-identified data from all of the billing entities with NHP claims in the state over 5 years so that we could assess differences in rates of screening, positive screening, and mental health treatment by site as well as gender and age. In a recent paper (Murphy et al., 2020), we analyzed data on all 72,842 WCVs from the 304 sites with at least 30 claims in a single year (2016).

2017 Submission
Data Source 2:
In samples drawn from the total MassHealth claims database, 2000 visits per year of patients in 3 cohorts (years 2007, 2010, & 2012), chart reviews by nurses who were experienced medical record reviewers to obtain more detailed information about demographics, pediatric visits, claims, and mental health outcomes. The data we present was a secondary analysis of existing data by researchers from the University of Massachusetts Medical School of information they had collected and analyzed for a state contracted evaluation of the CBHI program.

Data Source 3:
3 years of WCV and BH screening data for a data set of 10,334 4–17-year-old children with Medicaid from four Massachusetts General Hospital-affiliated outpatient clinics.
Although data sources 1 and 2 had provided information on many WCV and screening variables and permitted analysis by region and age group, they did not include data on a clinic-by-clinic basis. To obtain an initial estimate of whether it was possible to identify differential rates of BH screening at the clinic level we turned to our own hospital system where we were able to obtain IRB approval to obtain billing data for psychosocial screens for a relatively large sample of well-child visits in four MGH affiliated outpatient pediatric clinics from 7/1/2014-12/31/2016. Using the same CPT billing code for screening (96110) used by the state and focusing only on children with Medicaid health insurance, we obtained WCV, BH screening, and demographic data on 35,608 visits. Because of data inconsistencies due to the hospital's switch to Epic in 2016, in this proposal we focused on the 6,801 cases screened in only 2016.

2021 Submission
Data source 4:
The denominator was all children ages 5.5 to 17.99 years old who were seen for a WCV in the 15 pediatric practice networks between January 1, 2018, and December 31, 2019. For each of the patients, specific fields in the practice network's Electronic Data Warehouse (EDW), which are allocated for PSC item level data and computed total and subscale scores, were searched for evidence that a PSC was completed in conjunction with that visit. Investigators were also able to access some demographic data on each patient. With IRB approval, we received a limited use dataset of all variables and cases and published the results in a recent paper (Murphy et al., 2020b).
[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]
2017 Submission
Data source 1: Medicaid claims data for all 6-month to 20-year-old patients (and, alternatively, for just the 3–17-year-old patients) seen for well child visits in the state of Massachusetts from 1/1/2008-3/31/2017

2021 Submission
Data source 1x: Medicaid claims data for WCV from 1/1/2016 - 12/31/2016 for children enrolled in the state's largest Medicaid HMO, a subset of Data Source 1.

2017 Submission
Data source 2: Random samples of 2000 visits per cohort drawn from Data source 1 from 3 cohorts: 7/1/2006-12/31/2009; 7/1/2009-12/31/2010; 7/1/2011-12/31/2012

Data source 3: Billing and demographic data extracted from Partners Healthcare System data warehouse for claims for WCV and BH screens and demographic data for all 4–17-year-old patients from four MGH-affiliated outpatient pediatric clinics from 7/1/2014-12/31/2016

2021 Submission
Data source 4: Psychosocial screening and demographic data were extracted from the EDW for all patients 5.5 to 17.99-years of age seen for WCV from 1/1/2018 - 12/31/2019.
[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]
Facility
Health Plan
Population: Regional and State
[Response Ends]

2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.

[Response Begins]
2017 Submission
Data source 1: The base sample is claims data from the entire state of Massachusetts for whom WCV claims for youth 6 months to 20 years of age were submitted. These data include all clinics and practices in the state. The CBHI website also breaks the data out according to the six regions of the state and six age groups (<6-months, 6-month to 2- years-old, 3- to 6-years-old, 7- to 12-years-old, 13- to 17-years-old, and 18- to 20-years-old).

2021 Submission
Data source 1x: This was a subsample of Data Source 1 but in this case, actual claims data for all WCV for members of the state's largest Medicaid HMO (Neighborhood Health Plan). The measured entities were all 664 practices/providers who submitted claims for WCV. The analytic sample was the 304 clinicians or practices which billed for at least 30 WCV in 2016.

2017 Submission
Data source 2: This 6,000 case sample was drawn from Data Source 1, with stratified random samples of 2000 cases per cohort for the 2007, 2010, and 2012 cohorts, with 500 cases each for 4 age groups (6 months to 2 years, 3 to 5 years; 6-11 years, and 12-20).

Data source 3: Billing and demographic data extracted from Partners Healthcare System data warehouse for claims for WCV and BH screens and demographic data for all 4-17 year old patients from four MGH-affiliated outpatient pediatric clinics from 7/1/2014-12/31/2016.

2021 Submission
We tested reliability and validity of the CPT billing codes in the claims data using chart reviews in a randomly selected subsample of cases from the three facilities combined.

Data source 4: The measured entities were a group of 15 independently owned outpatient pediatric practices from eastern Massachusetts. Called the Affiliated Pediatric Practices (APP) because of their affiliation with the Massachusetts General Hospital/Brigham Hospital network, these practices care for about 100,000 children, about 80% of whom have commercial health insurance (about 20% have Medicaid). The PSC is the required psychosocial screen at all WCV, and all data are collected electronically via iPads in the waiting room or over the internet, with item level data filed in specific fields of the EMR. Reliability and validity of the EHR coding were established through chart review of a randomly selected subsample of all cases.
[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]
Data source 1:

Summary data from the CBHI Behavioral Health Screening Cumulative Quarterly Report; posted on BHSCQR website

Measurement Period dates of service from 1/1/2008 to 3/31/2017

Statewide data for all claims for all well child visits for all Medicaid insured children ages .5 -20 years of age: N=4,721,790

Numerator (screens with visit): 2,965,923

Statewide average: 62.8%

Minimum: 14.2%

Maximum: 71.9%

Standard Deviation: 12.4%

95% Confidence interval: 58.8% to 66.8%

The CBHI BHSCQR also presents the same statewide Medicaid screening data broken down for each of the state's regions for each of the 37 calendar quarters from Q1 (January through March of 2008) to Q37 (January through March of 2017) of CBHI. Table 1 (below) reports the rates of screening in the state's six regions with data from the four quarters in each year averaged for simpler presentation in this proposal. Because the BHSCQR website does not break out the data by age group, the information in the table below is for the full sample of all ages, but the trends should be large enough that that the patterns shown below should be quite similar to those for just the 3–17-year-olds. As the table shows, the range in rates of screening vary widely, from a low of 21.55 in the Boston region in 2008 to a high of 85.27 in the Western region in first quarter of 2017. The percentages of well child visits with screens goes up substantially in all regions over the first few years of the initiative but after that the rank ordering remains relatively consistent across regions. Western Massachusetts always has the highest rate of screening, Metro West and Boston the lowest and Northeast, Southeast, and Central Massachusetts in the middle.

Table 1. Rates of Screening in different regions of Massachusetts from 2013-2017 by Region

Region	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017 ***
Western	38.14%	66.13%	75.34%	79.16%	79.79%	80.33%	81.62%	83.76%	81.27%	85.27%
Central	32.61%	51.99%	60.61%	63.22%	67.23%	70.59%	71.09%	70.90%	72.33%	72.31%
Northeast	32.70%	56.36%	60.93%	60.31%	55.35%	65.08%	67.16%	68.43%	71.46%	67.38%
Metro West	25.89%	47.45%	51.92%	51.80%	57.72%	60.47%	59.83%	54.90%	49.84%	47.72%
Southeast	36.88%	61.89%	70.57%	75.94%	78.18%	80.19%	80.66%	77.87%	75.91%	73.95%
Boston	21.55%	45.91%	55.28%	57.26%	60.53%	64.49%	67.57%	65.12%	63.27%	58.76%

Note. Rates of screening were calculated by the year from 2008-2016 and for the first quarter of 2017

*Data available only for the 1st quarter of 2017

Table 2 below presents the distribution of rates of screening in greater detail, taken from this time directly from BHSCQR website for all 222 measurement points (37 quarters in all six regions), illustrating even more dramatically the wide range of rates of screening across the state and over nearly a decade.

Table 2 The distribution of screening rates by decile broken down by quarters for each region from January 2008 through March 2017 (6 regions x 4 quarters x 9.25 years) for children of all ages

0-9%	1
10-19%	5
20-29%	4
30-39%	8
40-49%	23
50-59%	40
60-69%	63
70-79%	58
80-89%	20
90-100%	
Total (Regions*Quarters*Years)	222
Min:	8.21%
Max:	85.65%
Range:	77.44%
Median:	64.53%

Data Source 1b: Statewide data broken down by for **children ages 3-17 years**

Denominator /Well child visits for all children 3 -17 years of age: 2,361,475

Numerator (screens with visits): 1,681,764

Statewide average: 71.2%

Minimum: 39.98%

Maximum: 79.14%

Standard Deviation: 11.3%

95% Confidence Interval: 64.2% to 78.2%

We included data source 1b focusing just on the 3-17 year old children in the BHSCQR website data as this is the age group that was screened with the PSC, and prior studies (Hacker et al., 2016; Savageau et al., 2016) have shown that the PSC was the measure used for 67% of the children in the CBHI 3-17 year old age group. By multiplying the total number of screens in this age group by 67%, we can estimate that approximately 1,126,782 PSC's were administered over the first 9.25 years of CBHI.

Table 3 (below) shows the Massachusetts statewide number of well child visits, number of screens, and percent of visits with screens, for just the 3–17-year-old (PSC screened) children with Medicaid from January 2008 (start of CBHI) to March of 2017. Data in this table are taken directly from the CBHI BHSCQR but with the four quarters of each year aggregated together so that the totals for each year could be seen more clearly. As the table shows, the rate of screening rose from approximately 39.98% for its first year (2008) to 65.72% for its second year to over 70% for its third year, and then remaining in the 70% range in all of the six years since. Not shown in this table but present in the data shown on the CBHI BHSCQR are the figures for the first quarter of 2008 (which show a base rate of 17.8% during the first three months of the program). Although not posted on the CBHI BHSCQRbut reported by two different groups (Hacker et al., 2016; Savageau et al., 2016) with access to claims data for 2007 is that the rate of formal screening during the year prior to the start of CBHI was less than 5%. A rate of formal screening that started and then remained at less than 5% was also reported for the state of California (which had no requirement for the use of formal screens) for 2008 and 2009 when the rate in Massachusetts had climbed to about 65%.

Table 3. Statewide rates of formal psychosocial screening for 3–17-year-olds during WCV from 1/1/2007 to 3/31/2017

Year	Denominator (Total Visits)	Numerator (Total with Screens)	% Visits with Screens
12/31/07-12/31/2008	202,376	80,910	39.98%
1/1/2009-12/31/2009	219548	144276	65.72%
1/1/2010-12/31/2010	234823	171114	72.87%
1/1/2011-12/31/2011	243209	179305	73.72%
1/1/2012-12/31/2012	252357	190286	75.40%
1/1/2013-12/31/2013	265826	208549	78.45%
1/1/2014-12/31/2014	287690	228467	79.14%
1/1/2015-12/31/2015	290313	217282	74.84%
1/1/2016-12/31/2016	300307	215611	71.80%
1/1/2017-3/31/2017	65,026	45,964	70.69%
Total	2,361,475	1,681,764	71.22%

*Cell intentionally left blank

2021 Submission
Data Source 1x

One of the most important criteria for a quality measure is the ability to identify differences in performance across sites. Because we did not have access to actual claims from CBHI in Data Sources 1 and 2 at the time of our 2017 submission, we could not assess site differences in screening in a large sample. Instead, we turned to our own hospital system and were able to obtain billing data for pediatric psychosocial screens for a relatively large sample of well-child visits for only four outpatient pediatric clinics as described below with regard to Data Source 3.

Several years later, we were finally able to get access to individual encounter data from the largest Medicaid HMO in Massachusetts, a subsample of the CBHI data described as Data Source 1.

The study sample (Murphy et al., 2020b) included 76,022 WCVs from 664 sites in 2016. For statistical reasons, we focused our analysis on the 304 sites that completed at least 30 WCVs in that year. This included most of the cases (N=72,842 WCVs). A mean of 240 WCVs (median = 125), with a range of 30-1,880 WCVs, were completed per site. Overall screening rates ranged from 0.0-100.0%, with a mean site screening rate of 75.4%. A majority of sites (n=197; 73.9%) had a screening rate above 70.0%. A majority (76.1%) of patients were below the age of 13.00 and 51.0% were male (Table 1). The ICC for screening rates or the proportion of total variance in screening rates accounted for by site variation in the 304 sites was 0.38 (95% confidence interval [CI]=0.25- 0.50). Given that this proportion is ≥10.0%, the observed between-site screening rate differences can be considered meaningful (Fleiss & Jack 1979; Zou & Donner, 2004). The percent of cases that screened positive was 7.5% overall, with a range of 0-100%. The ICC for positive screening rates in the 245 sites that also had screening results available was 0.10 (95% CI=0.00-0.29). The proportion of total variance in screening rates accounted for by site variation was 10.0%, just within the lower bound of what is considered a meaningful between-site variation.

Relationship Between Site-Level BH Screening and BH Service Use

To explore the relationship between site-level overall screening, positive screening, and BH service utilization rates in the six months following the BH screen, we first ran a correlation between site-level BH screening rates and rates of subsequent BH service use. A positive but statistically non-significant relationship between a site's BH screening rate and the rate of subsequent BH service use was found (Pearson's r=0.08, p=0.17). However, when overall site screening rates were dichotomized to contrast sites falling below (n=108 sites) with those falling above (n=196 sites) the mean screening rate of 71.9%, (as determined by CBHI) and when subsequent BH service use was dichotomized as falling below (n=45 sites) or above (n=259 sites) 6.0% (the mean rate of BH treatment in CBHI), sites that screened above 71.9% were significantly more likely to have BH treatment rates of 6.0% or above when compared to sites with less than 71.9% screening (88.3% vs. 79.4%; ⍰ 2=4.34; df=1, p<.05), suggesting a positive relationship between BH screening rates and rates of subsequent BH service use at the site-level (Table 2).

Statewide data on all 4-17 year old children enrolled in NHP in 2016

Denominator / WCVs for all of these children aged 4-17 years = **72,842**

Numerator (visits with screens): **55,759 with completed screens (76.5%)**

Average across sites in 2016: **71.2%**

Standard Deviation: **31.3%**

Minimum for a practice: **0%**

Maximum for a practice: **100%**

ICC: 0.38

95% Confidence Interval for ICC: **0.25-0.50**

2017 Submission
Data Source 2: CBHI Cohort Data for chart review sample

Measurement Period dates of service in 2007, 2010, and 2012

Denominator /Well child visits for all children: 4,977

Numerator (screen at WCV): ~1700

Average: 51.7%

Min: 1.5%

Max: 88.9%

Standard Deviation: 35.1%

95% Confidence interval: 35.5% to 87.2%

This second data source is a ~ 6000 visits subsample of the CBHI statewide data retrieved from chart reviews supplemented by administrative claims data.

Differences in rates of screening in three cohorts of about 2000 cases of pediatric outpatients (age 4-17 years) seen before (2007) and after (2010, 2012) the start of CBHI. As shown in Table 4, consistent with the statewide CBHI data, the Southeast and Western regions consistently demonstrated significantly higher rates of screening, the Northeast and Metro West demonstrated the lowest rates of screening, and Central Massachusetts and Boston demonstrated screening rates in the middle.

Table 4. Rates of Formal Screening for each year for children ages 4-17

	Formal Screen in 2007		Formal Screen in 2010		Formal Screen in 2012	
	No	Yes	No	Yes	No	Yes
Region		*[1]		***[2]		***[3]
Western	139 (95.2%)	7 (4.8%)	22 (11.9%)	163 (88.1%)	33 (15.2%)	184 (84.8%)
Central	64 (94.1%)	4 (5.9%)	30 (23.3%)	99 (76.7%)	31 (23.9%)	99 (76.1%)
Northeast	125 (91.2%)	12 (8.8%)	63 (37.5%)	105 (62.5%)	71 (39.0%)	111 (61.0%)
Metro West	84 (94.4%)	5 (5.6%)	54 (32.5%)	112 (67.5%)	62 (37.4%)	104 (62.7%)
Southeastern	197 (98.5%)	3 (1.5%)	52 (23.0%)	174 (77.0%)	27 (11.1%)	217 (88.9%)
Boston	149 (98.0%)	3 (2.0%)	51 (25.4%)	150 (74.6%)	33 (18.2%)	148 (81.8%)

*Cell intentionally left blank

2007: $^{1}x^{2}=13.33$ $p=.021$ **2010**: $^{2}x^{2}=36.33$ $p<.001$ **2012:** $^{3}x^{2}=75.21$ $p<.001$

Data Source 3**:** Medicaid screening data from four Massachusetts General Hospital outpatient clinics

Measurement Period dates of service from 7/1/2014 to 12/31/2016

Denominator /Well child visits for all children 4-17 years of age: 10,334

Numerator (screen at WCV): 7,915

Average: 76.6%

Min: 9.4%

Max: 91.7%

Standard Deviation: 38.5%

95% Confidence interval: 38.8% to 100.0%

One of the most important criteria for a quality measure is differences in performance across sites. Although the data from the CBHI BHSCQR summarized above provide strong evidence that there are consistent differences in screening rates for different regions of the state and different age groups, our lack of access to the actual data made it impossible to explore differences between different pediatric practices, one of the intended uses of this type of quality measurement. Since it was not possible for us to obtain the actual data from the state of Massachusetts and since the sample from Data Source 2 was too small to permit analyses by clinic, we turned to our own hospital system and were able to obtain billing data for pediatric psychosocial screens for a relatively large sample of well child visits in four MGH-affiliated outpatient pediatric clinics. Using the same CPT billing code for screening (96110) used by the state and focusing only on children with Medicaid health insurance, we obtained data on 10,827 children aged 4-17 who had at least one pediatric well child visit from July 2014 through December 2016 in one of the four clinics.

Differences in rates by clinic: As shown in Table 5 below, the differences in rates of mental health screening at well-child visits were substantial and statistically significant between these four clinics. In 2016, for example, rates ranged from 0% to 88.1% (p < .001) to in the four clinics. Over the three years of data, the rank orders of screening rates among the four clinics were relatively constant with Clinic A and B as always the highest and C the lowest in each of the three years (not shown).

Table 5: Rates of screening in four MGH affiliated outpatient pediatric clinics in 2016

			2016 (n=6,801 children): January 1, 2016 – December 31 2016				
Variable	**Level**	**Overall***	**Clinic A**	**Clinic B**	**Clinic C**	**Clinic D**	**p-value over 4 sites**
Billed for MH screen	**Yes**	69.7% (4743)	86.1% (3422)	88.1% (1215)	0.0% (0)	24.1% (106)	<.001

<u>2021 Submission</u>
Data Source 4

Data for children ages **5.5-17.99** years

Denominator of children who had well-child visits

In Phase 1 (pre-QI implementation; Jan 1, 2018 - April 31, 2018): **9,212**

In Phase 2 (QI implementation; May 1, 2018-December, 2018): **26,025**

In Phase 3 (post-QI implementation; 2019): **40,880**

Numerator (visits with screens):

Overall screening rate across 18 practices

In Phase 1: **26.1%** (range for the 18 practices: **14.2% - 52.4%**)

In Phase 2: **89.3%** (range for the 18 practices: **70.2%-96.0%**)

In Phase 3: **77.9%** (range for the 18 practices: **52.4%-84.8%**)

Denominator of children who had a well-child visit in either Phase 2 or 3: **45,949**

Numerator (visits with screens): **39,202** (out of the 45,949 visits in either or both years) or **85.3%**

Denominator of children who had a well-child visit in **both** Phase 2 & 3: **20,956**

Numerator (visits with screens): **16,223** (out of the 20,956 in both years) or **77.4%**

The study sample (Murphy et al., 2020a) included all 76,117 children aged 5.5 to 17.99 years who were seen for WCV over a two-year period in a network of 15 pediatric practices. The analytic sample was 20,956 patients with PSC screens in both years. With a total of 77.4% of all patients completing a PSC screen, this study demonstrated the feasibility of screening: 1) with the PSC specifically (in the CBHI samples, about 1/3 of screens used were other measures); 2) using a fully electronic system (completed by parents online from home or on iPads in the waiting room, which then scored and displayed in the EHR instantaneously); and 3) in a predominantly commercially-insured sample. Rates of screening and risk differed significantly by insurance type, sex, age, and practice, although in the ICC analyses, the differences in rates of screening and risk across 15 sites did not reach statistical significance.

[Response Ends]

2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.

[Response Begins]
2021 Submission
In our 2017 submission, Data Sources 1-3 were all based on administrative claims data and reliability and validity were tested using those data compared to chart review. The Massachusetts statewide CBHI program requires all children seen for WCV to have a psychosocial screen administered and (for accountability purposes) for providers to bill for that screen using CPT code 96110 . Providers are also required to use a CPT modifier code to indicate whether the screen is positive or negative. In our 2017 submission, the validity of using the presence or absence of CPT code 96110 to verify the presence of a PSC was checked using a chart review as the gold standard. Having research assistants review the WCV progress notes in the EHR's of 45 cases established the validity of using CPT code 96110 as the measure of whether a PSC had been given. Reliability for this type of coding was established by having two different research assistants code a subsample of these cases independently.

In our 2021 submission, we did not re-do the billing code review since validity had already been established for Data Sources 1-3 using this method. However, since Data Source 4 involved a different kind of data, we conducted another chart review using it as the gold standard for our calculations. First of all, since Data Source 4 was based on a pediatric sample that was primarily commercially insured and in which providers are not required to bill for their screens, billing could not be used as an accurate reflection of whether screening had been a part of a WCV. In the network we studied for Data Source 4, the standard was to administer all screens electronically using iPads from the waiting room or over the internet from home with item level data automatically and instantaneously scored and stored in the network's enterprise data warehouse (EDW). Data Source 4 was established by downloading the data found in the PSC item and subscale fields for all patients with WCV in the sample. To establish the validity of this PSC data, a review of the WCV progress note was again used as the gold standard. Random samples of all cases were reviewed to establish validity by ascertaining whether the data from the EDW that was provided to researchers was confirmed in the chart review. Reliability was established by having two coders blindly and independently rate a subsample of these cases.
[Response Ends]

2a.08. List the social risk factors that were available and analyzed.

For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g., census tract), or patient community characteristics (e.g., percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

[Response Begins]
2021 Submission
Data Source 2: Researchers who had published a study (Savageau et al., 2016) based on chart reviews of ~6,000 WCVs to evaluate psychosocial screening among Medicaid-insured children in Massachusetts were able to reopen their datasets to assess the impact of several social risk factors on psychosocial screening (Savageau et al., 2017). In their 2016 paper, the researchers reported on a major increase in the use of formal screening tools (with state approved measures like the Pediatric Symptom Checklist) from 2007 (<5%) to 2010 and 2012 (>70%). Using the reopened data sets, these patterns were found to be very similar for children from all racial, ethnic, and language groups. Rates of positive screening did not differ according to race or ethnicity but were slightly but significantly lower for children from non-English speaking families after the start of the screening program. The impact of these risk factors on subsequent service use was negligible: for children who screened positive, rates of BH service use in the six months after the screen did not differ significantly for any of the above three groups.
[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter “see validity testing section of data elements”; and enter “N/A” for 2a.09 and 2a.10.

Reliability Testing

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.
[Response Begins]
Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)
[Response Ends]

2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.

Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.

[Response Begins]
2017 Submission
Although not relevant to the PSC as a process measure, it may be important to note that the reliability and validity of the PSC as a tool have been established over the years in many studies (Bergmann et al., 2018; Lavigne, Meyers, & Feldan, 2016; Pourat et al., 2017). The reliability/validity of using the CPT code in administrative data to signify the completion of an actual PSC screen has not been reported in a published paper

2021 Submission
but was described in our 2017 NQF submission. In our 2017 NQF submission, **validity** was established by comparing billing codes indicating that a screen had been given with a review of the progress note for the WCV to look for evidence of the screen. Inter-coder **reliability** was assessed by comparing the chart review results of two raters who examined the progress note blindly with regard to each other’s coding. What follows is the description of this reliability coding that was presented in section 2a2.3 of the Measure Testing Form of that submission.

"As noted above, we assessed the inter-rater reliability of this medical record coding by having a second research assistant code 30 of the 90 charts (10 from each clinic) selected at random. For the 96110-by-chart-review coding, the ratings of the two RA's agreed on the presence or absence of a screen in 28/30 (93%) of the charts and kappa was .86. In the same 30-case subsample, the second coder also found evidence of a well child visit on all of the dates billed (100% agreement with billing codes and with the first coder). The two coders agreed on the name of the specific screen used in 94% of the cases."

In our 2021 submission, we again used chart reviews as the gold standard. This time, research assistants examined the WCV progress note and supporting documents to see whether (1) the item and subscale scores downloaded from the EDW were also found in the progress note and (2) whether WCV that did not have item- and subscale-level data for the WCV in the EDW download were also bereft of such data in the patient's EHR.

[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, NQF Measure Evaluation Criteria).

[Response Begins]
2021 Submission
In our 2017 submission, using chart review as the gold standard for validity, the 96110 code in claims data had a sensitivity of 86.2% and a specificity of 100.0% and an overall rate of 91% intercoder agreement in coding the presence of a behavioral health screen during a well child visit. Kappa was .84.

In our 2021 submission, two coders reviewed twenty charts selected at random to assess whether the EDW record that a PSC given could be verified in a chart review. Inter-coder reliability was 100%.
[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]
2021 Submission
For both the 2017 (using billing codes) and 2021 (using EDW data) NQF submissions, intercoder agreement was over 90%. Levels of agreement this high are considered to indicate an acceptably high level of reliability and therefore both methods of assessing the completion of a PSC screen as a part of a WCV appear to be reliable.
[Response Ends]

Validity Testing

2b.01. Select the level of validity testing that was conducted.

[Response Begins]
Patient or Encounter-Level (data element validity must address ALL critical data elements)

Systematic assessment of face validity of performance measure score as an indicator of quality or resource use (i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance)
[Response Ends]

2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

[Response Begins]
2021 Submission
I. Encounter level data

Data Source 3: Because the validity of using the CPT code in billing data to indicate that a PSC had been completed had never been assessed, in our 2017 NQF submission we conducted a cross-validation analysis in the one sample for which we had access to individual case data, to assess whether the billing data was a valid indicator of whether a psychosocial screen was administered at a patient's well-child visit. To do this, the presence and/or absence of a CPT billing code (96110) associated with each patient's well child visit in 2016 was collected, and then each patient's chart information in the electronic medical record was evaluated to ascertain the presence or absence of a psychosocial screen by either the mention of a behavioral health screen in the progress notes or a copy of the PSC in the patient's chart on the same date of the well-child visit associated with the presence or absence of the 96110 code. What follows is what we submitted in section 2a2.2 of the Measure Testing Form. In the paragraph that follows we have substituted the word 'validity' for the word 'reliability' to distinguish more clearly and consistently between the two.

Reliability and validity of the PSC as a process measure
2021 Submission [Data Source 1; 2017]

2017 Submission
To test the validity of the coding of the administration of a PSC tool, we performed chart reviews to ascertain whether the CPT code (96110) used to bill for screening corresponded with evidence in the progress note for the WCV that a PSC or other approved screen had actually been given. Using the presence/absence of evidence of PSC administration in the visit note as the gold standard, we calculated the sensitivity, specificity, and kappa of the CPT code charge for the screen. In order to test the validity of other critical data elements, we also coded: 1) whether there was a progress note or other evidence of an encounter in the chart documenting that a WCV had occurred on the date it was billed and 2) whether the code used for the WCV was correct based on the age of the child. In addition, as already described above, we evaluated the inter-rater reliability of all of these assessments by having a second-rater code (blind to first reviewer's coding) one third of the cases on all three variables.

2a2.3. For each level of testing checked above, what were the statistical results from validity testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of* validity *statistics from a signal-to-noise analysis*)

Validity testing for CPT billing codes compared to chart review as gold standard

From the dataset of all WCV with 4-17-year-old children covered by Medicaid from the four MGH-affiliated outpatient pediatric clinics for calendar year 2016, we were able to access complete EMR data from three clinics resulting in a sample of 6,462 children. For each of these, the dataset included the presence or absence of the CPT billing code used for psychosocial screening (96110) on the same day as the well child visit, age and several other demographic and billing variables. In each clinic, 30 WCV were selected and a research assistant opened the electronic medical record for each patient, checking to see whether the progress note for the visit documented that a CBHI approved psychosocial screen had been given, either by the mention of a behavioral health screen and a score, or by finding a PDF of the PSC or another screen in the patient's chart. In each clinic, we coded the first 15 and the last 15 WCV of 2016. Cases marked as restricted access (highly confidential) were skipped and replaced with the next case that matched the selection criteria.

Out of the 90-case sample, 52 (57.8%) of the cases had a 96110-billing code on the date of their well-child visit in 2016 and 38 (42.2%) of the cases did not have a 96110 code. Using documentation of a screen in the chart as the gold standard, there were 59 cases in which we found documentation of a screen in the chart. The 96110 coding correctly identified 52 of these patients (88.1%) as having had a screen and of the 31 patients with no screen in their charts, all 31 (100.0%) were correctly identified as not screened (no 96110 code). Kappa was .84.

The Kappa statistic has the following interpretation:

0.00=Poor; 0.01 – 0.20=Slight; 0.21 – 0.40=Fair; 0.41 – 0.60=Moderate;

0.61 – 0.80= Substantial; 0.81 – 0.99=Almost perfect agreement

Therefore, in this sample, using a billing code to establish the presence of a psychosocial screen during a WCV was found to have a very high level of validity.

We also checked for other key data elements using the same method. The first coder also assessed whether there was documentation in the chart that a WCV had occurred on the date billed and found that this coding was perfect: all 90 cases (100%) had documentation that a WCV had been completed on the date that it was billed. Since there are 10 different CPT codes for WCV depending on the age of the child, we also checked 30 cases to compare the age of the child with the specific WCV CPT code used and found that the age appropriate WCV CPT code had been used in 100% of the cases.

To evaluate previous reports (Romano-Clarke et al., 2014; Savageau et al., 2017; Savageau et al., 2016) that about 40% of all screens done in CBHI (and 2/3 to ¾ of the screens for 4–17-year-olds) were PSC's, we checked the name of the screen listed in the visit note. In the 90 case chart review sample, the first coder found that of the 59 cases in which there was documentation of a screen, 54 (92%) reported that the PSC was the screen that had been given, 3 (5.1%) mentioned another CBHI approved screen (the Parents Evaluation of Developmental Status (PEDS)) and 2 (3.4%) did not give the name of the screen. Since the PEDS can be used for children from 6 months to 8 years of age, the choice of the screen in these three cases was appropriate and since the PSC has been well-validated for children who are 4-17 years old and since the coder found that all charts that mentioned a specific screen were of children who were within this age range, this coding also showed that the choice of instruments was age appropriate in 100% of cases.

2a2.4 What is your interpretation of the results in terms of demonstrating validity *(what do the results mean and what are the norms for the test conducted?*)

2017 Submission

This coding showed that that the presence of the PSC or other brief BH screen during WCV could be validly coded from the presence of the CPT 96110 code in administrative claims data. This approach to coding was also validated by the finding that 100% of the WCV that were billed for were documented by EMR notes from the same day and that the age codes for the WCV were congruent with the age of each child. ...The presence of the PSC as the named screen in 92% of these cases confirmed previous reports (Romano-Clarke et al., 2014, Savageau et al., 2016) that the great majority of the screens used with 4–17-year old's in the CBHI were PSC's.

2021 Submission

Data Source 4: This data set used encounter-level data to document the presence of in the EHR of data from an electronically completed PSC-Tool as a part of each patient's WCV. Validity of the electronic coding was defined as finding item level data in the WCV progress note for cases coded as having had a completed psychosocial screen and also NOT having a scanned screen or a physician's report of a screen in the WCV progress note or in the documents section of the EHR for cases in which the EDW data did not show item level data for the PSC. Research staff reviewed random samples of WCV in which the EDW had found or not found data from the PSC. This method tested the validity of the electronic version of the PSC as a process measure.

Eighty cases (60 EDW negative and 20 EDW positive) were randomly selected for chart review. In 15 of the sixty EDW negative cases, evidence of a PSC screen was found in the progress note or in a scanned form, despite its absence from the EDW. In all 20 of the cases in which the EDW had provided PSC data, the coders found PSC data in the patient's progress note. Thus, overall sensitivity of the EDW coding was 60%, the specificity was 75%, and overall agreement between EDW data and chart review was 81.3% (65/80).

II. Systematic assessment of face validity of performance measure score as an indicator of quality or resource use

In one of the new sources of data we were able to explore one hypothesized validator of pediatric psychosocial screening as an indicator of quality: *whether practices that had higher rates of psychosocial screening had significantly higher rates of mental health service use*. In Data Source 1x we began by showing that using intra-class correlation, the rates of screening and positive screening differed significantly among the 304 screening entities.

The ICC for screening rates (the proportion of total variance in screening rates accounted for by site variation) was 0.38 (95% confidence interval [CI]=0.25- 0.50). Given that this proportion is ≥10.0%, the observed between-site screening rate differences can be considered meaningful. The percent of cases that screened positive was 7.5% overall, with a range of 0-100%. The ICC for positive screening rates in the 245 sites that also had screening results available was 0.10 (95% CI=0.00-0.29). The proportion of total variance in screening rates accounted for by site variation was 10.0%, just within the lower bound of what is considered a meaningful between-site variation (Murphy et al 2020b).

We next explored the relationship between site-level overall screening, positive screening, and BH service utilization rates in the six months following the BH screen. We ran a correlation between site-level BH screening rates and rates of subsequent BH service use. A positive but statistically non-significant relationship between a site’s BH screening rate and the rate of subsequent BH service use was found (Pearson’s r=0.08, p=0.17). However, when overall site screening rates were dichotomized to contrast sites falling below (n=108 sites) with those falling above (n=196 sites) the mean screening rate of 71.9%, (as determined by CBHI) and when subsequent BH service use was dichotomized as falling below (n=45 sites) or above (n=259 sites) 6.0% (the mean rate of BH treatment in CBHI), sites that screened above 71.9% were significantly more likely to have BH treatment rates of 6.0% or above when compared to sites with less than 71.9% screening (88.3% vs. 79.4%; ⍰ 2=4.34; df=1, p<.05), suggesting a positive relationship between BH screening rates and rates of subsequent BH service use at the site-level (Table 2).

It may be important to note at this point that the relationship between increases in psychosocial screening and increase in outpatient mental health services as the key validator of the value of routine screening has been established much more strongly in a recent study with a much larger sample (N=~10,000,000). Hacker and colleagues (2016) used the Medicaid claims data from Massachusetts described in our Data Source 1 to compare changes in mental health service use over the first 18 months of the CBHI program in that state with any possible changes in mental health services for children seen for WCV in California over the same period of time when rates of psychosocial screening did not change. The authors concluded that, "The goal of BH screening is to identify previously unidentified children with BH issues and provide earlier treatment options. The short-term outcomes of the Massachusetts policy suggest that screening at preventive care visits led to more BH-related outpatient visits among vulnerable children." Although our results in the Data Source 1x study failed to reach statistical significance with regard to the relationship between screening and increased service use, the fact that they were in the predicted direction in our Data Source 1x and that they were significant in the larger Hacker et al, sample, provides face validity for the claim that implementing routine psychosocial screening in pediatrics is associated with improved outcomes: a larger percent of vulnerable children receiving outpatient mental health services.

For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

2021 Submission
For the encounter level data described above, the accuracy of the billing code data and the EDW data were compared to chart reviews using reliability and validity statistics. For system level data, the validity of psychosocial screening as an indicator of quality was established by showing that outpatient behavioral health service use increased in the state that adopted routine screening but not in a state that did not increase screening.
[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]
2021 Submission
In the claims data analyses, the sensitivity of the coding vs chart review was 86% and the specificity was 100% with an overall level of agreement of 91%. In the EDW analyses, the sensitivity of the electronic PSC coding vs chart review was 48.8% and the specificity was 100% with 73.8% overall agreement.
[Response Ends]

2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)

[Response Begins]
2017 Submission
Although based on a very small sample, the validity of the 96110-coding appeared to be high.

2021 Submission
The validity of the electronic form coding was only moderate. Data Source 4 probably underestimates the accuracy of electronic coding since the pediatricians in the APP network had been screening routinely with paper forms for several years prior to adopting electronic screening and the chart review showed that a number of the clinicians still used paper forms for the parent- or youth-reported PSC 35 which had not been programmed for electronic data entry. In any case, the average of 74.4% SN + SP represents an acceptable level of validity for a network that had just adopted electronic administration of the PSC. Going forward, even in a system used to screening with paper forms, it should be relatively easy to emphasize more strongly the need to enter paper-based scores by hand or to accept only electronically completed forms during the in-service training for pediatricians. It seems likely that this would lead to greater screening accuracy.
[Response Ends]

2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.

Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]
2021 Submission
Statewide screening data from Massachusetts between 2008-2017 had shown large and persistent differences in rates of screening by region and age group. And when intra-class correlation analysis (ICC) was used to investigate clinic-level differences in screening rates within a large, statewide sample of pediatric primary care practices (Data Source 3), statistically significant differences in rates of screening and rates of positive screening were found. The

clinical significance of positive screening was evident in Hacker's (2014b) study of the CBHI in Massachusetts, in which children without a BH history who screened positive were almost four times more likely to receive BH services and even children who did have BH histories were more than twice as likely to get subsequent BH services. Even in a relatively homogeneous, relatively small sample of patients in a network of 15 pediatric practices serving primarily commercially insured children, higher rates of psychosocial screening were significantly associated with higher rates of subsequent mental health service utilization.
[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

[Response Begins]
In the comparisons between Massachusetts regions, one of the six had surpassed the 80% benchmark sought by the state, two others were in the low to mid 70% range which was close to the mark. Two other regions were in the high 40% range, illustrating the large differences and large room for improvement in some regions. In the small sample of MGH clinics there was one clinic that was not screening at all and another that was screening at a very low rate.

The study (Hacker et al, 2016) that found significant differences in rates of screening with formal measures in pediatrics in Medicaid samples in California vs those in Massachusetts provides strong empirical evidence of the relationship between statistically and clinically meaningful differences in rates of screening and their relationship to increased behavioral health service use in a statewide comparison.

2021 Submission
Hacker et al 2016 summarized their methods:
"We calculated monthly population utilization rates and adjusted for managed care enrollment, eligibility status (such as foster care and poverty), race, and Hispanic ethnicity as described above. The numerator for rates was the presence of a claim for service in the calendar month for one of the diagnostic or procedure codes listed above. The denominator for rates was all youths eligible to receive services (in other words, they were currently insured by Medicaid) and of screen able age in the same calendar month. That is, we did not remove youths from the denominator in subsequent study months when they received BH screening in prior study months. We further adjusted the population utilization rates for seasonality using the Census Bureau algorithm (proc X-11 in SAS) (22,23). Interrupted Time-Series Analysis We fit segmented regression models (24–26) for each of the BH utilization rates in the 24 months prior to mandatory screening and 18 months post mandate. The segmented regression models included terms for the change in intercept (immediate-level change), secular trend (overall slope), and change in trend (increase versus decrease in slope postimplementation). We then constructed difference-indifferences models comparing Massachusetts with California to determine whether there were significant changes in outcomes related to the policy."

Later in this paper they summarized their results*:*

"Figure 2 shows the monthly rate of any BH-related outpatient service utilization. The adjusted rate was about 35 per 1,000 youths per month in the two years prior to the mandate, with a slightly increasing secular trend (.8 per 1,000, p=.03). After the phase-in period for screening, rates of outpatient use increased dramatically in the fall of 2008, to about 50 per 1,000 youths per month. More specifically, it began to rise during the phase-in period and then remained stable thereafter. The trend in utilization increased at about 4.5 per 1,000 youths per month (p,.001). No increase was seen in California. Estimates for policy (immediate change), time (secular trend), and time after the mandate (post mandate trend) resulting from the regression models are shown in Table 2.

And as noted above, ICC analyses in a statewide sample of 304 sites showed statistically significant differences across sites in rates of screening and positive screening, with some comparisons of rates of service utilization showing significantly higher rates of BH service use in sites that had higher rates of screening..
[Response Ends]

2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]
2021 Submission
Statistically and clinically significant differences in rates of screening and in outpatient mental health service use in relation to increased screening were consistently found in large samples like those mentioned in 2b.05 and 2b.06. Although they were not consistently found in smaller samples, it seems reasonable to conclude for now that the findings from the larger samples provide the most accurate picture, namely that large scale psychosocial screening programs are associated with increases in outpatient mental health services. This conclusion is supported by the recently reaffirmed recommendation for routine psychosocial screening as a part of all WCV by the American Academy of Pediatrics Task Force on Mental Health (AAP, 2020).
[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]
Missing data, in this case PSC screening scores is not an issue since those patients who have WCV and are not assessed in the measurement period remain in the denominator.
[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]
Patients who are seen and not assessed with the PSC during the assessment period are included in the denominator of patients with eligible WCVs.
[Response Ends]

2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.

In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.

[Response Begins]
Missing data is not an issue for this measure as constructed.
[Response Ends]

Note: This item is directed to measures that are risk-adjusted (with or without social risk factors) OR to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eCQMs). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]
No, there is only one set of specifications for this measure
[Response Ends]

2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.

Describe the steps—do not just name a method. Indicate what statistical analysis was used.

[Response Begins]
[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]
[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]
[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]
N/A or no exclusions
[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?

[Response Begins]
N/A
[Response Ends]

2b.17. Provide the statistical results from testing exclusions.

Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.

[Response Begins]
N/A
[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.

[Response Begins]
N/A
[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]
No risk adjustment or stratification
[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]
N/A
[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]
N/A
[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]
Published literature
Internal data analysis
[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]
N/A
[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]
N/A
[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]
N/A
[Response Ends]

2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]
N/A
[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]
N/A
[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]
N/A
[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]
N/A
[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]
N/A
[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]
N/A
[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]
N/A
[Response Ends]

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]
Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.
[Response Begins]
Some data elements are in defined fields in electronic sources
[Response Ends]

3.03. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.

[Response Begins]
2021 Submission
In medical systems that do not utilize electronic data elements for PSC score capture, progress notes in electronic or paper medical records can be reviewed for evidence that a screen was administered. Paper or stand-alone electronic records of screens can be compared to daily logs of completed WCV to provide evidence of percent of cases that were screened. It is not uncommon for simpler approaches like these to be implemented initially in order to get patients and providers used to screening. Once the culture of a practice has evolved to expect routine screening, a switch to a more automated electronic system can be implemented.
[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]
None as of this writing.
[Response Ends]

3.06. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

[Response Begins]
2021 Submission

Although our Data Source 4 and studies in progress in two other pediatric practice networks have provided very strong evidence that electronic approaches to screening are feasible at network scale for almost 90% of all WCV in the first year and still almost 80% in the second, our validity assessment for this method of screening showed only moderate accuracy. Although all PSC's found in the EDW were also found in the charts, the converse was not always true. We found that it was not uncommon for providers to administer paper versions of the PSC (e.g., for non-English speaking parents, the PSC was administered but in a language that had not been programmed for electronic administration) and since these forms would have to be hand entered into the template if they were to show up in the EDW, they often were not, and thus were not counted when only the EDW data download was reviewed.

This was not an insurmountable problem, since provisions can be made to increase the number of PSC forms in other common languages that are made available electronically and/or to train medical assistants to hand enter the items for PSC's completed in other languages (as is already done in several large clinics).

Viewed as a work in progress, current data on the relative accuracy of electronic PSC administration as well as anecdotal evidence of the adoption of routine screening with the PSC in at least a dozen large pediatric networks around the US (and in other countries) provide strong evidence that routine psychosocial screening in pediatrics continues to become more and more common.
[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

3.07. Detail any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm),

Attach the fee schedule here, if applicable.

[Response Begins]
No proprietary elements are used in implementing this measure. There are no fees, licenses or other requirements needed to use any aspect of the measure or the instrument.
[Response Ends]

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement, in addition to demonstrating performance improvement.

4a.01.

Check all current uses. For each current use checked, please provide:

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]
2017 Submission
Payment Program
Professional Certification or Recognition Program
Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

2021 Submission
The Affiliated Pediatric Practices (APP) network offered all of its 50+ providers the chance to receive American Board of Pediatrics MOC credit for participating in a QI project in which each clinician assessed her/his rate of screening with the PSC successively over a three month period and tried to improve it.

In the Reliant Medical Group (RMG), another multi-practice network in Massachusetts, eighteen pediatricians participated in a similar ABP MOC project with the PSC, this time using it as a first stage screen for depression in adolescents.

In both the APP and RMG MOC projects, higher and higher benchmarks for rates of screening were set for each subsequent month of the project.

In the Commonwealth of Massachusetts, the Childrens' Behavioral Health Initiative (CBHI) requires that all pediatric providers who bill for a WCV must have used one of about twelve approved measures to screen for psychosocial problems as a part of that WCV. Providers are also required to **bill for administering the screen** using a CPT code and to indicate whether the screen was positive or negative by using a CPT modifier code. If these conditions are met, the provider (or more commonly, the practice) receives a payment over and above the reimbursement for the WCV. In Data Source 4, participating providers received small incentive payments if they meet benchmark goals for an assortment of QI goals each quarter, including percentage of WCV with completed screens.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]
Professional Certification or Recognition Program
Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
Quality Improvement (internal to the specific organization)
Measure Currently in Use
[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]
N/A
[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.

[Response Begins]
N/A
[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail how many and which types of measured entities and/or others were included. If only a sample of measured entities were included, describe the full population and how the sample was selected.

[Response Begins]
2021 Submission
In the first years of the CBHI program, the performance results based on screening rates for psychosocial and medical problems or interventions for each clinic in the state were used as the basis for an in-person QI review by the Massachusetts state Medicaid program. Results for rates of psychosocial screening and dozens of other quality measures are still calculated bi-annually and sent to administrators at each site. In the early years of the program, all of the state's Medicaid HMO's were also required to report on their rates of psychosocial screening.

According to the PCC Plan Provider Handbook (http://www.mass.gov/eohhs/docs/masshealth/provider-services/forms/pcc-handbook.pdf) “The PCC Profile Report provides information on selected clinical measures, such as pediatric behavioral health, well-child care, and women’s cancer screening, that may be used to improve health care delivery and, ultimately, the health outcomes of PCC Plan members. Most of the profile measures display rates of performance for a PCC’s practice as well as rates for each service location, if applicable, and for the PCC Plan as a whole. Prior rates for PCC practices and the PCC Plan are also presented to show trended rates for these clinical indicators. Summary data are provided to help with the identification of barriers to care.” (page 24). For sites with at least 180 PCC Plan members… a Regional Network Managers (RNMs) … visits in order to review with the PCC the PCC Profile Report, the PCC Reminder Report, and the PCC Care Monitoring Registries. By reviewing the reports and discussing how rates reflect the PCC practice’s performance, RNMs assist practices to identify areas for improvement and to develop action plans to improve performance and the delivery of high-quality health care to members”. (page 23).

In the networks that have adopted an electronic approach to psychosocial screening (Murphy et al 2020a, Murphy et al, 2016), reports which include item and subscale data, and interpretation of risk scores are now available to pediatricians instantaneously, as soon as the PSC is completed. This means that the assessment of each patient's psychosocial functioning can easily be incorporated into the pediatrician's overall evaluation, planning, and recommendations (Arauz Boudreau et al, 2019). In many if not most cases, results of PSC screening will also be shared with the parent who completed it.
[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]
2021 Submission
As noted above, in sites that implement psychosocial screening using electronic methods, results are available for clinicians and patients instantaneously in the EHR. Since calculating the PSC total score requires only simple addition of the weighted item scores, even clinicians who administer the PSC in paper and pencil format can ascertain whether the screen is positive or negative takes in just a matter of seconds and since the cut off scores for risk are fixed (e.g., 28+ for the PSC 35), it is easy for clinicians to arrive at the assessment of overall psychosocial risk quickly and easily after scoring the form by hand at the start of a WCV. When practices implement routine psychosocial screening, they often begin with in-service trainings that familiarize clinicians with the process of screening and the meaning of screening results (Murphy, et al, 2020a).
[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]
2021 Submission
Although specific feedback has not been requested from sites that have implemented routine psychosocial screening with the PSC, anecdotal evidence suggests that an increasing number of pediatric practices and networks have implemented screening. Although the PSC is available to all without cost, most sites that implement screening write to us requesting formal permission to do so and we have received such requests from hundreds of sites (several each week for many years). Some support for this interpretation is provided by the fact that the number of papers published on the PSC has continued to grow, year over year, for example with more than 100 papers published since our 2017 NQF endorsement request.
[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]
Based on the above, the feedback has been overwhelmingly positive.
[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]
Not known
[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]
Not known
[Response Ends]

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

[Response Begins]
2021 Submission
Table 3 (section 1b.2) shows the increase in rates of screening in the full sample of 3–17-year-olds over the first 9.25 years of CBHI. Table 1 shows that these increases were present in all regions of the state. Rates of screening for all ages increased dramatically, 14-fold, from the period pre-CBHI.

Progress on improvement in psychosocial screening with the PSC-Tool over the past few years is evident in the increasing number of networks that have adopted routine psychosocial screening as the standard of care. One example of this is the shift from our previous NQF submission which focused on data based primarily on children with Medicaid in a state which made screening a requirement for them, to our new focus on Data Source 4 which describes a fully electronic system that has made it possible to administer, score, and store screens for commercially insured children without regard to billing for the screens. Now more than halfway through its fourth year and administering more than 30,000 PSC's a year, the system described in Data Source 4 has demonstrated the feasibility of screening for large general, commercially insured pediatric populations in real world settings. As noted above, we are working on data from two other large networks which now screen routinely with the PSC at all WCV and although we have not seen any data from them yet, we have heard that nearly one dozen other large outpatient pediatric networks have implemented such systems. The country of Peru has made routine psychosocial screening a required part of its annual pediatric care for all of the adolescents in that country and has specified the PSC as the measure that must be used. Plans are underway to mandate screening for younger children as well.
[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]
There have been no reports of unintended negative consequences to individuals or populations.
[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]
2017 Submission
Benefits include:
-Increasing widespread use of a simple but effective PRO tool that can be used for screening, diagnosis, and the monitoring of treatment outcomes for psychosocial problems (Massachusetts, California, and other states).
-Increased national use of the measure (PSC is being used in the SAMHSA National System of Care Expansion Evaluation and in the state of California child mental health outcomes assessment program), Mental Health America making the PSC and PSC-Y available for free and to tens of thousands of youth or their parents.

2021 Submission
With the increasing number of networks requiring PSC screening in pediatric settings, the technology for electronic administration has improved (there are now a great many software systems ranging from simple overlays to full featured EHR's) that make it relatively easy to electronically screen patients routinely ,at less expense, and this, in turn has probably led to an increase in the use of the PSC as a mandatory tool in behavioral health networks as a way to meet the QI goal of using standardized measures in outcomes based care. The state of California's requirement that the PSC be administered at intake for mental health services and then every six months and the Community Behavioral Health Association of Maryland's similar requirement in its 72 member agencies are examples of this.
[Response Ends]

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02 if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)
[Response Begins]
2801: Use of First-Line Psychosocial Care for Children and Adolescents on Antipsychotics
1365: Child and Adolescent Major Depressive Disorder (MDD): Suicide Risk Assessment
0576: Follow-Up After Hospitalization for Mental Illness (FUH)
0108: Follow-Up Care for Children Prescribed ADHD Medication (ADD)
0711: Depression Remission at Six Months
1885: Depression Response at Twelve Months- Progress Towards Remission
0712: Depression Assessment with PHQ-9/ PHQ-9M
1884: Depression Response at Six Months- Progress Towards Remission
0710: Depression Remission at Twelve Months
[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)
[Response Begins]
[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]
2021 Submission
Before discussing two non-NQF endorsed measures, we thought it was important to begin with some context. Outside of the PSC (NQF#3332), a review of all NQF measures found only nine other NQF-endorsed measures that related to pediatric mental health. Four of these did not overlap with the PSC at all: 108: Follow-up care for children prescribed ADHD medication; 576: Follow-up after hospitalization for mental illness; 1365: Suicide risk assessment; and 2801: Use of First-Line Psychosocial Care for Children and Adolescents on Antipsychotics.

The remaining five NQF-endorsed pediatric mental health measures do relate to the PSC because all involve depression, which is one the several areas relevant to PSC screening. The five measures are: 0710: Depression Remission at Twelve Months; 0711: Depression Remission at Six Months; 0712: Depression Assessment with the PHQ-9/PHQ-9M; 1884: Depression Response at Six Months-Progress Towards Remission; 1885: Depression Response at Twelve Months- Progress Towards Remission.

The five NQF-endorsed adolescent depression measures all rely on the PHQ-9. This questionnaire screens for depression only and has been validated for use only with adolescents (not younger children). The PSC does not compete with these five adolescent depression measures because the PSC does not have the same target population and does have a much broader focus. All five of the currently endorsed measures that use the PHQ-9 apply only to adolescents already diagnosed with depression. The reason for using the PHQ-9 is to have a benchmark measure of whether, on re-screening, the adolescents initially diagnosed with depression have responded or remitted over six or twelve months. In contrast, the target population for the PSC is children as well as adolescents and it includes 100% of both children and adolescents, not just the 5-10% of adolescents who are depressed.

Another use of the PHQ-9 which was more closely related to the PSC is no longer endorsed by NQF but is still endorsed by other standard-setting agencies like the American Academy of Pediatrics, HEDIS, and the CHIPRA Core Set is NQF measure #418/418e, Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan. Minnesota Community Measurement is the steward for these related measures.

As mentioned above, in addition to the PHQ-9, there are two non-NQF endorsed instruments that are related to the PSC and that should be discussed as possible competing measures: the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Child Behavior Checklist (CBCL; Achenbach, 1991). Robert Goodman is the author of the SDQ, and Thomas Achenbach is the steward of the CBCL. Both of these measures are frequently used in research and in mental health (but not pediatric) clinical settings. Before reviewing them below, it may be important to note that there are also several other measures that have been endorsed as mental health screens for younger children. But these do not compete with the PSC because, for the most part, they are used to screen children aged five or younger.
[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]
No
[Response Ends]

5.05. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

[Response Begins]
N/A No endorsed competing measures to harmonize.
[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]
2021 Submission
In our 2017 NQF proposal, we gave the following reasons that routine screening in pediatrics with the PSC is superior to routine screening with the PHQ-9:

2017 Submission
The age range for the PHQ-9 (NQF 712) has recently been expanded to include youth 12 to 17 years of age with a diagnosis of depression. The currently submitted measure, the PSC, screens for a broader band of problems (other emotional problems like anxiety as well as other types of problems like attention and behavior) and a larger age

range (3-17) than the PHQ-9. Along with the PHQ-9, the PSC is actually one of the specific tools mentioned by the US Preventive Services Task Force as a screen for depression in youth (Forman-Hoffman et al., 2016). Although studies have shown that the PSC identifies about 80% of the youth with depression who are found with the PHQ-9, only about half of the youth with serious psychosocial problems on the PSC are identified with the PHQ-9 (Richardson et al., 2010).The PSC is a representative of a broader class of screening tools (brief broadband psychosocial screens) that are required for use in conjunction with pediatric well child visits in the Massachusetts EPSDT program. Other similar broadband tools are the Strengths and Difficulties Questionnaires and the Child Behavior Checklist. The Massachusetts EPSDT CBHI program provides a short (now 13) list of approved tools (both broad and narrow band) and allows the pediatrician to use the one deemed most appropriate for each case. In a review of nearly 6000 medical charts, Savageau and her associates found that about 40% of all screens were PSC's compared to only about 1% that were PHQ-9's (Savageau et al., 2016; Savageau et al., 2017, May) suggesting that the PSC is at least in the past ten years more widely used by pediatricians in Massachusetts.

The reference list is included in the attached appendix.

2021 Submission

As noted above, there are two other potentially competing (but not NQF endorsed) measures. Like the PSC, both the CBCL and SDQ have been validated for use in identifying children and adolescents of school age (e.g., ages 4-17) who may have mental health risk. All three questionnaires include a core set of items that assess problems that are rated by a parent, a teacher, or the youth him/herself. Although the CBCL and SDQ could, theoretically, be used as routine screens in pediatrics, neither is widely used for this purpose. As an example, the SDQ --but not the CBCL—is currently endorsed by CBHI for use in psychosocial screening in pediatrics for children and adolescents with Medicaid in Massachusetts. The CBCL, although probably the most frequently used mental health instrument for children and adolescents in the world, with 120+ items and a 20+ minute administration time, is too long and time-consuming to meet the need for mental health screening in pediatrics. Although initially included as an endorsed measure for CBHI, the CBCL lost its CBHI endorsement after a year or two. The SDQ, although brief enough to use in pediatrics, is more complicated to administer and score (two different forms for primary vs secondary school age groups and more subscales to calculate and interpret). Although it is still a CBHI endorsed measure, as noted above, a review of six thousand charts from this initiative reported 0% use of the SDQ compared to about 40% use of the PSC and PSC-Y. The fact that the SDQ has recently joined the CBCL in requiring payment for each administration has also made its use in pediatrics even less likely (all PSCs and translations can be downloaded at no cost).

[Response Ends]

Appendix

Supplemental materials may be provided in an appendix.: Available in attached file
Attachment: 3332_Full NQF Appendix_11-2-21_posted 11-2-21.docx

NQF #3332 Psychosocial Screening Using the PSC-Tool

Appendices

Table of Contents

Children's Behavioral Health Initiative 91

Pediatric Symptom Checklist 95

1. Reference list for PSC papers since 2017 NQF submission 100

2. References for all papers cited in 2021 (and 2017) NQF submissions 110

Children’s Behavioral Health Initiative

Lucke, Savageau, Simons, Jellinek, Murphy (2017). Assessing Differences in Mental Health Screening by Region Before and After Implementation of an Innovative Statewide Program for Children with Medicaid. *Poster presentation at Massachusetts General Hospital Clinical Research Day*, Boston, Massachusetts; October 5, 2017.

CONTROL ID: 2779015
TITLE: Assessing Differences in Mental Health Screening by Region Before and After Implementation of an Innovative Statewide Program for Children with Medicaid
PRESENTATION TYPE: MGH Clinical Research Day Abstract
CURRENT CATEGORY: Individual Submission

AUTHORS (LAST NAME, FIRST NAME): Lucke, Cara[1]; Savageau, Judith A.[2]; Simons, Jack[3]; Jellinek, Michael[4]; Murphy, J. Michael[1,4]
INSTITUTIONS (ALL):
1. Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States.
2. Family Medicine and Community Health, University of Massachusetts Medical School, Worcester, MA, United States.
3. Executive Office of Health and Human Services, Commonwealth of Massachusetts, Boston, MA, United States.
4. Department of Psychiatry, Harvard Medical School, Boston, MA, United States.

KEYWORDS: Screening.

ABSTRACT BODY:

Introduction: Massachusetts implemented the Children’s Behavioral Health Initiative (CBHI) which contains mandates for routine pediatric behavioral health (BH) screening as well as outpatient services for children in need. Several studies have documented substantial increases in both screening and outpatient BH services over the eight years of the program consistently across race, ethnicity, and language. Because prior CBHI research had already documented a lack of racial/ethnic disparities in rates of screening and rates of positive screening, the current study explored potential differences in rates of screening and rates of positive screening among children residing in different geographic regions of Massachusetts.

Methods: We assessed differences in rates of screening and of positive screening for children (4-17 years) in three cohorts of about 2000 cases of pediatric outpatients seen before (2007) and after (2010, 2012) the start of CBHI. Data were from chart reviews supplemented by administrative claims data.

Results: The current study confirmed previously published reports of a major increase in the use of formal screening tools from 2007 (<5%) to 2010 and 2012 (>70%). Although there were increases in rates of screening in all regions from 2007 to 2010 and 2012, there were consistent significant differences in these rates. The Western and Southeastern regions demonstrated the highest rates of screening in both 2010 and 2012 (~77-89%) while the Northeast and Metro West regions demonstrated the lowest rates of screening (~61-68%). There were also consistent significant differences in rates of positive screenings in both 2010 and 2012, with the Central and Boston regions having the lowest rates of positive screens (~4-8%) and the Northeast and Metro West regions having the highest rates of positive screens (~23-33%).

Conclusion: The current study demonstrated that previously reported increases in formal screening with the advent of the CBHI applied to children and adolescents regardless of geographic region in Massachusetts. Although all geographic regions showed increases in formal screening after CBHI, there were nevertheless significant and persistent differences in rates of screening and positive screens by region. Interestingly, the Northeast and Metro West regions demonstrated significantly lower rates of

screening but significantly higher rates of positive screens among those who were screened. Further work to assess whether these differences in screening are associated with differences in service utilization and health outcomes is warranted.
Department: Psychiatry
Consortium affiliations or banner authorships: (none)

Savageau, Simmons, Lucke, Jellinek, Murphy (2017). Assessing Disparities in Mental Health Screening and Services Before and After Implementation of an Innovative Statewide Program for Children with Medicaid. *Paper presented at Pediatric Academic Society,* San Francisco, CA; May 7, 2017.

CONTROL ID: 2696718
TITLE: Assessing Disparities in Mental Health Screening and Services Before and After Implementation of an Innovative Statewide Program for Children with Medicaid
ROLE TYPE: Abstract
CURRENT CATEGORY: General Pediatrics
CURRENT SUBCATEGORY: Health Services Research
KEYWORDS:
AUTHORS (FIRST NAME, LAST NAME): Judith Savageau[1], Jack Simmons,[2] Cara Lucke,[3] Michael Jellinek,[4] J. Michael Murphy[3]
INSTITUTIONS (ALL)
1. Family Medicine and Community Health, University of Massachusetts Medical School, Worcester, MA, United States.
2. Executive Office of Health and Human Services, Commonwealth of Massachusetts, Boston, MA, United States.
3. Massachusetts General Hospital, Boston, MA, United States.
4. Harvard Medical School, Boston, MA, United States.
TITLE: Assessing Disparities in Mental Health Screening and Services Before and After Implementation of an Innovative Statewide Program for Children with Medicaid

Background: Massachusetts recently implemented the Children's Behavioral Health Initiative (CBHI) which contains mandates for routine behavioral health (BH) screening in pediatrics as well as outpatient services for children in need. Although several studies have documented substantial increases in both screening and outpatient BH services over the 8 years of the program, until now no studies have assessed these patterns for children from minority vs. non-minority groups.
Design/Methods: We assessed the presence of racial, ethnic, and language disparities in rates of screening, positive screening, and receipt of subsequent BH services for children (4-17 years) who screened positive in three cohorts of about 2000 cases of pediatric outpatients seen before (2007) and after (2010, 2012) the start of CBHI. Data were from chart reviews supplemented by administrative data.
Results: There was a major increase in the use of formal screening tools (with state approved measures like the Pediatric Symptom Checklist) from 2007 (<5%) to 2010 and 2012 (>70%). These patterns were very similar for children from all racial, ethnic, and language groups. Rates of positive screening did not differ according to race or ethnicity but were significantly lower for children from non-English speaking families after the start of CBHI. For children who screened positive, rates of BH service use in the six months after the screen did not differ significantly for any of the minority vs. non-minority groups.
Conclusion(s): The current study demonstrated that previously reported increases in formal screening with the advent of the CBHI applied to children from all three minority groups. Although there did appear to be some disparities for children from non-English speaking families, the fact that rates of positive screening and of service utilization after positive screening did not differ significantly for children from racial or ethnic minority groups suggested a predominant lack of disparities in CBHI. These positive findings were to some extent counterbalanced by the finding that only about 25% of positively screened children (minority and non-minority) received BH services in the six months after screening. This in turn suggested a persisting unmet need for services and the need for further research identifying possible barriers to service.
(no table selected)
(No Image Selected)

Content Type Expertise: Health Services/Clinical Outcomes Research, Patient Safety/Quality of Care/Quality Improvement, Policy/Advocacy/Legislation, Population Health/Global/Social Determinants of Health
APA SIG Comm Region: None of These
First Author Trainee? No, Not a Trainee

Pediatric Symptom Checklist

Pediatric Symptom Checklist Website
https://www.massgeneral.org/pediatric-symptom-checklist/
Child's Name _________________
Today's Date _________________
Date of Birth _________________
Record Number ___________________
Filled out by _____________

Pediatric Symptom Checklist (PSC-35)—Parent Report

Emotional and physical health go together in children. Because parents are often the first to notice a problem with their child's behavior, emotions, or learning, you may help your child get the best care possible by answering these questions. Please mark under the heading that best fits your child.

		Never (0)	Sometimes (1)	Often (2)
1. Complains of aches/pains	1	_____	_____	_____
2. Spends more time alone	2	_____	_____	_____
3. Tires easily, has little energy	3	_____	_____	_____
4. Fidgety, unable to sit still	4	_____	_____	_____
5. Has trouble with a teacher	5	_____	_____	_____
6. Less interested in school	6	_____	_____	_____
7. Acts as if driven by a motor	7	_____	_____	_____
8. Daydreams too much	8	_____	_____	_____
9. Distracted easily	9	_____	_____	_____
10. Is afraid of new situations	10	_____	_____	_____
11. Feels sad, unhappy	11	_____	_____	_____
12. Is irritable, angry	12	_____	_____	_____
13. Feels hopeless	13	_____	_____	_____
14. Has trouble concentrating	14	_____	_____	_____
15. Less interest in friends	15	_____	_____	_____
16. Fights with others	16	_____	_____	_____
17. Absent from school	17	_____	_____	_____
18. School grades dropping	18	_____	_____	_____
19. Is down on him or herself	19	_____	_____	_____
20. Visits doctor with doctor finding nothing wrong	20	_____	_____	_____
21. Has trouble sleeping	21	_____	_____	_____
22. Worries a lot	22	_____	_____	_____
23. Wants to be with you more than before	23	_____	_____	_____
24. Feels he or she is bad	24	_____	_____	_____
25. Takes unnecessary risks	25	_____	_____	_____
26. Gets hurt frequently	26	_____	_____	_____
27. Seems to be having less fun	27	_____	_____	_____
28. Acts younger than children his or her age	28	_____	_____	_____
29. Does not listen to rules	29	_____	_____	_____
30. Does not show feelings	30	_____	_____	_____
31. Does not understand other people's feelings	31	_____	_____	_____

32. Teases others	32	______	______	______
33. Blames others for his or her troubles	33	______	______	______
34. Takes things that do not belong to him or her	34	______	______	______
35. Refuses to share	35	______	______	______

Total score ________________

Does your child have any emotional or behavioral problems for which she/he needs help? () No ( ) Yes
Are there any services that you would like your child to receive for these problems? () No ( ) Yes
If yes, what services? ____

Pediatric Symptom Checklist (PSC-17)—Parent Report

Please mark under the heading that best describes your child:

		Never (0)	Sometimes (1)	Often (2)
1. Feels sad, unhappy	1	______	______	______
2. Feels hopeless	2	______	______	______
3. Is down on self	3	______	______	______
4. Worries a lot	4	______	______	______
5. Seems to be having less fun	5	______	______	______
6. Fidgety, unable to sit still	6	______	______	______
7. Daydreams too much	7	______	______	______
8. Distracted easily	8	______	______	______
9. Has trouble concentrating	9	______	______	______
10. Acts as if driven by a motor	10	______	______	______
11. Fights with other children	11	______	______	______
12. Does not listen to rules	12	______	______	______
13. Does not understand other people’s feelings	13	______	______	______
14. Teases others	14	______	______	______
15. Blames others for his/her troubles	15	______	______	______
16. Refuses to share	16	______	______	______
17. Takes things that do not belong to him or her	17	______	______	______

Does your child have any emotional or behavioral problems for which she/he needs help? () No ( ) Yes

Pediatric Symptom Checklist (PSC-35)—Youth Report

Please mark under the heading that best fits you:

		Never (0)	Sometimes (1)	Often (2)
1. Complain of aches or pains	1	______	______	______
2. Spend more time alone	2	______	______	______
3. Tire easily, little energy	3	______	______	______
4. Fidgety, unable to sit still	4	______	______	______
5. Have trouble with teacher	5	______	______	______
6. Less interested in school	6	______	______	______
7. Act as if driven by motor	7	______	______	______
8. Daydream too much	8	______	______	______
9. Distract easily	9	______	______	______
10. Are afraid of new situations	10	______	______	______
11. Feel sad, unhappy	11	______	______	______
12. Are irritable, angry	12	______	______	______
13. Feel hopeless	13	______	______	______
14. Have trouble concentrating	14	______	______	______
15. Less interested in friends	15	______	______	______
16. Fight with other children	16	______	______	______
17. Absent from school	17	______	______	______
18. School grades dropping	18	______	______	______
19. Down on yourself	19	______	______	______
20. Visit doctor with doctor finding nothing wrong	20	______	______	______
21. Have trouble sleeping	21	______	______	______
22. Worry a lot	22	______	______	______
23. Want to be with parent more than before	23	______	______	______
24. Feel that you are bad	24	______	______	______
25. Take unnecessary risks	25	______	______	______
26. Get hurt frequently	26	______	______	______
27. Seem to be having less fun	27	______	______	______
28. Act younger than children your age	28	______	______	______
29. Do not listen to rules	29	______	______	______
30. Do not show feelings	30	______	______	______
31. Do not understand other people’s feelings	31	______	______	______
32. Tease others	32	______	______	______
33. Blame others for his or her troubles	33	______	______	______
34. Take things that do not belong to him or her	34	______	______	______
35. Refuse to share	35	______	______	______

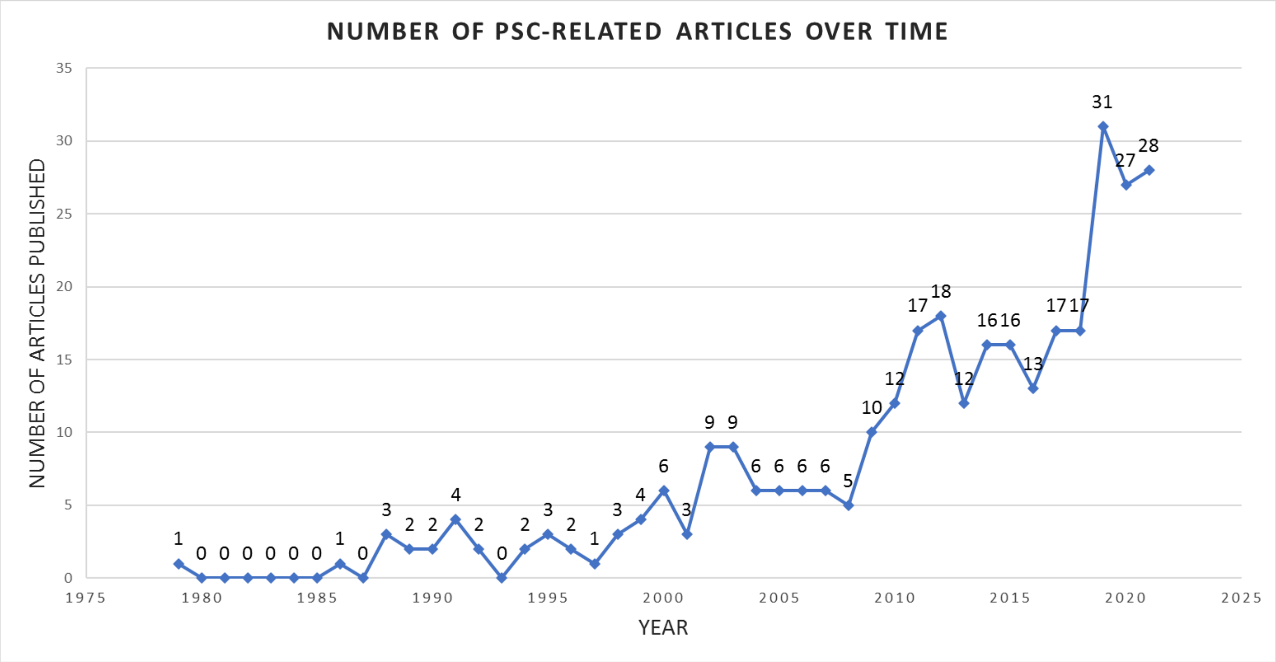
Pediatric Symptom Checklist (PSC-17)—Youth Report

Please mark under the heading that best fits you:

		Never (0)	Sometimes (1)	Often (2)
1. Fidgety, unable to sit still	1	______	______	______
2. Feel sad, unhappy	2	______	______	______
3. Daydream too much	3	______	______	______
4. Refuse to share	4	______	______	______
5. Do not understand other people’s feelings	5	______	______	______
6. Feel hopeless	6	______	______	______
7. Have trouble concentrating	7	______	______	______
8. Fight with other children	8	______	______	______
9. Down on yourself	9	______	______	______
10. Blame others for your troubles	10	______	______	______
11. Seem to be having less fun	11	______	______	_____
12. Do not listen to rules	12	______	______	______
13. Act as if driven by a motor	13	______	______	______
14. Tease others	14	______	______	______
15. Worry a lot	15	______	______	______
16. Take things that do not belong to you	16	______	______	______
17. Distract easily	17	______	______	______

1. Reference list for PSC papers since 2017 NQF submission

Figure 1: PSC Papers Over Time (1986-October 2021)



I. Empirical studies of PSC in Pediatric Mental Health Screening

○ A. Screening with the PSC is now the standard of care in many real-world settings

Arauz-Boudreau, A., Riobueno-Naylor, A., Haile, H., Holcomb, J.H., Lucke, C.M., Joseph, B., Jellinek, M.S., Murphy, J.M. (2019). How an electronic medical record system facilitates and demonstrates effective psychosocial screening in pediatric primary care. Clinical Pediatrics, 59(2). 154-162. doi: 10.1177/0009922819892038

Burke, R. E., Hoffman, N. D., Guy, L., Bailey, J., & Silver, E. J. Screening, Monitoring, and Referral to Treatment for Young Adolescents at an Urban School-Based Health Center. Journal of School Health, https://doi.org/https://doi.org/10.1111/josh.13089

Hansel T, Rohrer G, Osofsky J, Osofsky H, Arthur E & Barker C. (2017). Integration of Mental and Behavioral Health in Pediatric Health Care Clinics. Journal of Public Health Management & Practice, 23 Suppl 6 Supplement, Gulf Region Health Outreach Program, S19-S24. doi:10.1097/PHH.0000000000000649

Holcomb, J. M., Arauz Boudreau, A., Riobueno-Naylor, A., Haile, H., Lucke, C., Bergmann, P., Joseph, B. Jellinek, M., Murphy, J. M. (2020). Beyond initial screening: One-year follow-up of adolescents with internalizing problems on the Pediatric Symptom Checklist. Journal of Developmental and Behavioral Pediatrics."

Honigfeld, L., Macary, S. J., & Grasso, D. J. (2017). A Clinical Care Algorithmic Toolkit for Promoting Screening and Next-Level Assessment of Pediatric Depression and Anxiety in Primary Care. Journal of Pediatric Health Care, 31(3), e15-e23.

Jacobson, J. H., Pullmann, M. D., Parker, E. M., & Kerns, S. E. (2019). Measurement Based Care in Child Welfare-Involved Children and Youth: Reliability and Validity of the PSC-17. Child Psychiatry & Human Development, 50(2), 332-345.

Leiva, L., Antivilo-Bruna, A., Torres-Cortés, B., Peña, F., & Scquicciarini, A. M. (2021). Relationship between adverse childhood experiences and mental health: Implications for a nationwide school mental health program. Current Psychology. https://doi.org/10.1007/s12144-021-02342-x

McLaurin-Jiang, S., Cohen, G.M., Brown, C.L., Edwards, P., & Albertini, L.W. (2020). Integrated mental health training relates to pediatric residents' confidence with child mental health disorders. Academic Psychiatry, 44. 299-304.

Murphy, J. M., Nguyen, T., Lucke, C., Chiang, C., Plasencia, N., & Jellinek, M. (2018). Adolescent Self-Screening for Mental Health Problems; Demonstration of an Internet-Based Approach. Acad Pediatr, 18(1), 59-65. https://doi.org/10.1016/j.acap.2017.08.013

Murphy, J. M., Riobueno-Naylor, A., Haile, H., Holcomb, J. M., Dutta, A., Kandru, R., Shui, A. M., Lee, H., & Jellinek, M. (2020). Behavioral Health Screening and Service Use in a Statewide Sample of Medicaid-Eligible Pediatric Outpatients. Science Repository. https://doi.org/10.31487/j.PDR.2020.03.04

Murphy, J. M., Stepanian, S., Riobueno-Naylor, A., Holcomb, J. M., Haile, H., Dutta, A., Giuliano, C. P., Bernstein, S. C., Joseph, B., & Shui, A. M. (2021). Implementation of an electronic approach to psychosocial screening in a network of pediatric practices. Academic Pediatrics, 21(4), 702-709.

Shellman, A. B., Meinert, A. C., & Curtis, D. F. (2019). Physician Utilization of a Universal Psychosocial Screening Protocol in Pediatric Primary Care. Clinical Pediatrics, 58(9), 957–969. https://doi.org/10.1177/0009922819845878

Spencer, A.E., Baul, T.D., Sikov, J., Adams, W.G., Tripodis, Y., Buonocore, O., Jellinek. M., Murphy, J.M., & Garg, A. (2020). The relationship between social risks and the mental health of school-age children in primary care. Academic Pediatrics, 20(2). 208-215. doi: 10.1016/j.acap.2019.11.006

Young, N. D., & Takala, C. R. (2018). Sequential Screening to Improve Behavioral Health Needs Detection in Primary Care. Journal of the American Academy of Child & Adolescent Psychiatry, 57(8), 603–609. doi: 10.1016/j.jaac.2018.04.017

○ B. Additional PSC projects are demonstrating feasibility and outcomes of screening

Berger-Jenkins E, Monk C, D'Onfro K, Sultana M, Brandt L, Ankam J, Vazquez N, Lane M, Meyer D. Screening for Both Child Behavior and Social Determinants of Health in Pediatric Primary Care. J Dev Behav Pediatr. 2019 Jul/Aug;40(6):415-424. doi: 10.1097/DBP.0000000000000676
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 [Multicenter Study Randomized Controlled Trial Research Support, N.I.H., Extramural]. Clin Pediatr (Phila), 59(8), 787-800. https://doi.org/10.1177/0009922820920013
Muppidathi, S., Boj, J., & Kunjithapatham, M. (2017). Use of the pediatric symptom checklist to screen for behaviour problems in children. International Journal of Contemporary Pediatrics, 4(3), 886-889.
Omkarappa, D. B., Rentala, S., & Nattala, P. (2021). Psychiatric nurse delivered group-cognitive-behavioral therapy for internalizing behavior problems among children of parents with alcohol use disorders. Journal of Child and Adolescent Psychiatric Nursing, 1-9. doi: 10.1111/jcap.12329
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. doi:10.1007/s11121-016-0689-5
Stadnick, N. A., Martinez, K., Aarons, G. A., Lee, D. A., Van Cleave, J., & Brookman-Frazee, L. (2020). Pediatric Primary Care Perspectives on Integrated Mental Health Care for Autism. Academic Pediatrics, 20(8), 1140-1147. https://doi.org/https://doi.org/10.1016/j.acap.2020.03.006

○ C. Quality Assurance/Quality Improvement

Craig, R., Naganathan, S., Ballance, C., Schairer, J., & Fleur, R. S. (2018). Improving Rates and Referrals of Developmental and Mental Health Screening at Well Child Visits. In: Am Acad Pediatrics.
May, J., Kazee, N., Castillo, S., Bahroos, N., Kennedy, S., Castillo, A., & ... Van Voorhees, B. W. (2017). From silos to an innovative health care delivery and patient engagement model for children in Medicaid. Healthcare (Amsterdam, Netherlands), doi:10.1016/j.hjdsi.2016.12.008.
Riobueno-Naylor, A., Romo, S., Kazis, L., Wang, S., Lydon, M., Nelson, J., Fowler, L., Drexler, A., Kogosov, A., Haile, H., Ryan, C. M., Chang, P., Warner, P., Palmieri, T. L., Lee, A. F., Stoddard, F., Murphy, J. M., & Sheridan, R. L. (2019). Usefulness of a Novel System for Feedback of Patient-Reported Outcome Measures in Children Recovering From Burns, Journal of Burn Care & Research, https://doi.org/10.1093/jbcr/irz082

○ D. Reliability and validity of PSC

Ardıç, E., Ünsal, G., & Bayram, S. (2020). Validity and reliability of the pictorial pediatric symptom checklist. Journal of Psychiatric Nursing, 11(1). 20-27.
Bergmann, P., Lucke, C., Nguyen, T., Jellinek, M., & Murphy, J. M. (2018). Identification and utility of a short form of the Pediatric Symptom Checklist-Youth self-report (PSC-17-Y). European Journal of Psychological Assessment. https://phstwlp2.partners.org:3267/10.1027/1015-5759/a000486 (Supplemental)
Chaffin, M., Campbell, C., Whitworth, D. N., Gillaspy, S. R., Bard, D., Bonner, B. L., & Wolraich, M. L. (2017). Accuracy of a Pediatric Behavioral Health Screener to Detect Untreated Behavioral Health Problems in Primary Care Settings. Clinical Pediatrics, 56(5), 427-434. doi:10.1177/0009922816678412.
Chaurasiya, A., Ranjan, J. K., & Asthana, H. S. (2019). Hindi Adaptation of Pediatric Symptom Checklist-Youth Version (PSC-Y). Journal of Indian Association for Child & Adolescent Mental Health, 15(1), 69–84.
DiStefano, C., Barth, S. G., & Greer, F. (2019). Assessing Equivalency of PSC-17 Ratings: Does It Matter if Mixed or Grouped Item Format Is Used? Journal of Psychoeducational Assessment, 37(7), 920-924. http://dx.doi.org.treadwell.idm.oclc.org/10.1177/0734282918819566
DiStefano, C., Liu, J., & Burgess, Y. (2017). Investigating the Structure of the Pediatric Symptoms Checklist in the Preschool Setting. Journal Of Psychoeducational Assessment, 35(5), 494-505. doi:10.1177/0734282916647648.

Higuchi, T. I., Yuko; Uenishi, Hiroyuki; Yanagimoto, Yoshitoki; Ono, Mayuko; Ishida, Haruhiko; Kaneko, Kazunari. (2021). Reliability and Validity of the Japanese Version of the Youth-rating Pediatric Symptom Checklist Shortened Version. [Published in Japanese].
Leiva L, Rojas R, Peña F, Vargas B, María Scquicciarini A. Detectando las Dificultades Emocionales y Conductuales en la Escuela: Validación de PSC-17. Revista Iberoamericana de Diagncstico y Evaluacicn Psicolcgica. 2019;1(50):95-105. doi:10.21865/RIDEP50.1.08.
Liu, J., DiStefano, C., Burgess, Y., & Wang, J. (2020). Pediatric Symptom Checklist-17 Testing Measurement Invariance of a Higher-Order Factor Model Between Boys and Girls. European Journal of Psychological Assessment, 36(1), 77-83. https://doi.org/10.1027/1015-5759/a000495
Liu, J., Guo, S., Gao, R., & DiStefano, C. (2020). Investigating school children's behavioral and emotional problems using pediatric symptoms checklist-17 in a structural equation modeling framework. School Psychology International, 41(3). 257-275.
Öner, Ö., Karaca, S., & Cangür, Ş. (2019). The validity and reliability of pictorial pediatric symptom checklist in children of 4-5 ages. Anadolu Psikiyatri Derg, 20(1). 73-81. https://doi.org/10.5455/apd.39053
Parker, E. M., Jacobson, J., Pullmann, M. D., & Kerns, S. E. (2019). Identifying psychosocial problems among children and youth in the child welfare system using the PSC-17: exploring convergent and discriminant validity with multiple informants. Child Psychiatry & Human Development, 50(1), 108-120.
Studts, C. R., Polaha, J. & van Zyl, M. A. (2017). Identifying unbiased items for screening preschoolers for disruptive behavior problems. Journal of Pediatric Psychology, 42(), 476-486.

○ E. PSC as a proxy for MH in studies of chronic illness and other situations

Baugh, E. J. & Ballard, S. M. & Carter, L. & Tyndall, L. & Nolan, M. & Littlewood, K. (2019). The County-Wide Dissemination of Triple P Parenting Program: An Implementation Study. Progress in Community Health Partnerships: Research, Education, and Action 13(1), 73-81. Johns Hopkins University Press. Retrieved July 5, 2019, from Project MUSE database.
Fallucco, E. M., Wysocki, T., James, L., Kozikowski, C., Williams, A., & Gleason, M. M. (2017). The Brief Early Childhood Screening Assessment: Preliminary Validity in Pediatric Primary Care. Journal of Developmental & Behavioral Pediatrics, 38(2), 89-98.
Huang, K. Y., Bornheimer, L. A., Dankyi, E., & Aikins, A. D. G. (2018). Parental Wellbeing, Parenting and Child Development in Ghanaian Families with Young Children. Child Psychiatry & Human Development, 49(5), 833-841.
Jellinek, M., Bergmann, P., Holcomb, J. M., Riobueno-Naylor, A., Dutta, A., Haile, H., Sturner, R., Howard, B., & Murphy, J. M. (2021). Recognizing Adolescent Depression with Parent-and Youth-Report Screens in Pediatric Primary Care. The Journal of Pediatrics, 233, 220-226, e221.
Mancini, M. A. (2020). A Pilot Study Evaluating a School-Based, Trauma-Focused Intervention for Immigrant and Refugee Youth. Child and Adolescent Social Work Journal, 37(3), 287-300. https://doi.org/10.1007/s10560-019-00641-8
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. Fam Syst Health, 37(3), 244-248. https://doi.org/10.1037/fsh0000429
Sachser, C., Berliner, L., Holt, T., Jensen, T. K., Jungbluth, N., Risch, E., et al. (2017). International development and psychometric properties of the Child and Adolescent Trauma Screen (CATS). Journal of Affective Disorders, 210, 189-195. doi:10.1016/j.jad.2016.12.040.
Spencer, A. E., Chiang, C., Plasencia, N., Biederman, J., Sun, Y., Gebara, C., Center, M. G. H. C. H., Jellinek, M., Murphy, J. M., & Zima, B. T. (2019). Utilization of Child Psychiatry Consultation Embedded in Primary Care for an Urban, Latino Population. Journal of Health Care for the Poor & Underserved, 30(2), 637-652. https://doi.org/10.1353/hpu.2019.0047
Spencer, A. E., Plasencia, N., Sun, Y., Lucke, C., Haile, H., Cronin, R., ... & Biederman, J. (2018). Screening for Attention-Deficit/Hyperactivity Disorder and Comorbidities in a Diverse, Urban Primary Care Setting. Clinical pediatrics, 57(12), 1442-1452.

Tolliver, M., Polaha, J., Williams, S.L., & Studts, C.R. (2020). Evaluating the prevalence of child psychosocial concerns in rural primary care. Journal of Rural Mental Health, 44(2). 96-105.

Zalta, A. K., Bui, E., Karnik, N. S., Held, P., Laifer, L. M., Sager, J. C., & ... Ohye, B. (2017). Examining the relationship between parent and child psychopathology in treatment-seeking veterans. Child Psychiatry and Human Development, doi:10.1007/s10578-017-0743-y.

Use of the PSC to identify psychosocial problems in youth with chronic conditions

Asthma

Lee, C. C., Holder-Niles, F. F., Haynes, L., Chan Yuen, J., Rea, C. J., Conroy, K., ... & Bottino, C. J. (2019). Associations Between Patient-Reported Outcome Measures of Asthma Control and Psychosocial Symptoms. Clinical pediatrics, 58(3), 307-312.

McGovern, C.M., Arcoleo, K., & Melnyk, B. (2019). COPE for asthma: Outcomes of a cognitive-behavioral intervention for children with asthma and anxiety. School Psychology, 34(6). 665-676. doi: 10.1037/spq0000310

Pesek, R., Whitaker, B., Ararat, E., Sood, A., Pertzborn, M., Davidson, A., Cobb, K., Hines, K., Margiotta, C., Turner, J., Jefferson, A., & Perry, T. (2021). Pediatric Severe Asthmatics in a Southern US State Face Multiple Challenges and Barriers to Care [Meeting Abstract]. Journal of Allergy and Clinical Immunology, 147(2), AB46-AB46.

Burn Care

Riobueno-Naylor, A., Romo, S., Kazis, L., Wang, S., Lydon, M., Nelson, J., Fowler, L., Drexler, A., Kogosov, A., Haile, H., Ryan, C. M., Chang, P., Warner, P., Palmieri, T. L., Lee, A. F., Stoddard, F., Murphy, J. M., & Sheridan, R. L. (2019). Usefulness of a Novel System for Feedback of Patient-Reported Outcome Measures in Children Recovering From Burns, Journal of Burn Care & Research, https://doi.org/10.1093/jbcr/irz082

Riobueno-Naylor, A., Williamson, H., Canenguez, K., Kogosov, A., Drexler, A., Sadeq, F., DePamphilis, M., Holcomb, J., Stoddard, F.J., Lydon, M., Murphy, J.M., & Sheridan, R.L. (2021). Appearance concerns, psychosocial outcomes, and the feasibility of implementing an online intervention for adolescents receiving outpatient burn care. Journal of Burn Care & Research, 3;42(1):32-40. https://doi.org/10.1093/jbcr/iraa108

Childhood Cancer/Survivors

Groves, A. P., Gettinger, K., Druley, T. E., Kozel, B. A., Shinawi, M., Mohrmann, C., ... & Hayashi, R. J. (2019). Special Therapy and Psychosocial Needs Identified in a Multidisciplinary Cancer Predisposition Syndrome Clinic. Journal of pediatric hematology/oncology, 41(2), 133-136.

Michaud et al. (2020). Evaluating the Brief Parental Intake Form (BPIF) for psychosocial difficulties in childhood cancer survivors. Psycho-Oncology, 29. 1510-1512.

Celiac Disease

Chellan, D., Muktesh, G., Vaiphei, K., Berry, N., Dhaka, N., Sinha, S.K., Thapa, B.R., & Kochhar, R. (2019). Effect of gluten-free diet and compliance on quality of life in pediatric celiac disease patients. Journal of Gastroenterology and Hepatology Foundation, 3. 388-393. doi: 10.1002/jgh3.12172.

Diabetes

Semenkovich, K., Berlin, K. S., Ankney, R. L., Klages, K. L., Keenan, M. E., Rybak, T. M., ... & Eddington, A. (2019). Predictors of diabetic ketoacidosis hospitalizations and hemoglobin A1c among youth with Type 1 diabetes. Health Psychology, 38(7):577-585. doi: 10.1037/hea0000719.

Epilepsy

Afzal, E., Khan, W. I., Iqbal, M. K., Aftab, K., & Ahmad, T. (2021). Prevalence of psychosocial problems in children with and without epilepsy: A comparative cross-section study. The Professional Medical Journal, 28(04), 557-563.

Choudhary, S., Niranjan, N., Khichar, S., Berwal, P. K., & Barath, A. S. (2017). Behavioral Problems and Intelligence Quotient Changes in Pediatric Epilepsy: A Case-Control Study. J Neurosci Rural Pract, 8(4), 617-621. https://doi.org/10.4103/jnrp.jnrp_57_17

Functional Constipation

El-Sonbaty, M.M., Fathy, A., Aljohani, A., & Fathy, A. (2019). Assessment of behavioural disorders in children with functional constipation. Open Access Macedonian Journal of Medical Sciences, 7(23). 4019-4022.

Klages, K. L., Berlin, K. S., Silverman, A. H., Mugie, S., Di Lorenzo, C., Nurko, S., Ponnambalam, A., Sanghavi, R., & Sood, M. R. (2017). Empirically Derived Patterns of Pain, Stooling, and Incontinence and Their Relations to Health-Related Quality of Life Among Youth With Chronic Constipation. J Pediatr Psychol, 42(3), 325-334. https://doi.org/10.1093/jpepsy/jsw068

Hematopoietic cell transplantation

Buchbinder, D., Patel, S.K., Casillas, J.N., Nugent, D.L., Neudorf, S., Sender, L.S., Torno, L., Huszti, H., Zeltzer, L.K., & Parsons, S.K. (2019). Parent proxy assessment of sibling quality of life following pediatric hematopoietic cell transplantation. Health and Quality of Life Outcomes, 17(162).

Hip Preservation Surgery

Richard, H. M., Cerza, S. P., De La Rocha, A., & Podeszwa, D. A. (2020). Preoperative mental health status is a significant predictor of postoperative outcomes in adolescents treated with hip preservation surgery. Journal of Childrens Orthopaedics, 14(4), 259-265. https://doi.org/10.1302/1863-2548.14.200013

HIV

Joshi, D., Tiwari, M. K., Kannan, V., Dalal, S. S., & Mathai, S. S. (2017). Emotional and behavioral disturbances in school going HIV positive children attending HIV clinic. Medical Journal Armed Forces India, 73(1), 18-22.

Kefale, D., Boka, A., Mengstu, Z., Belayneh, Z., & Zeleke, S. (2019). Emotional and behavioral problems and associated factors among children and adolescents on highly active anti-retroviral therapy in public hospitals of West Gojjam zone, Amhara regional state of Ethiopia, 2018: A cross-sectional study. BMC pediatrics, 19(1), 141.

Hunger, Poverty, Victimization, and other Adverse Childhood Experiences

Burdzovic Andreas, J. B., & O'Farrell, T. J. (2017). Psychosocial problems in children of women entering substance use disorder treatment: A longitudinal study. Addictive Behaviors, 65, 193-197.

Choi, K.R., Stewart, T., Fein, E., McCreary, M., Kenan, K.N., Davies, J.D., Naureckas, S., & Zima, B.T. (2020). The impact of attachment-disrupting adverse childhood experiences on child behavioral health. The Journal of Pediatrics, 221. 224-229. https://doi.org/10.1016/j.jpeds.2020.03.006

Fonseca, A., Kral, M. C., Lally, M. D., Boan, A. D., & Macias, M. M. (2019). Bullying and ostracism in youth with and without ADHD: implications for risk and resilience. Children's Health Care, 48(2), 164-175.

Gaete, J. (2021). Validation of the Paediatric Symptom Checklist to Identify and Assess Psychosocial Problems among Early Adolescents in Chile, 2017. UK Data Archive. https://doi.org/http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.5255/UKDA-SN-854843

Hawkins, J. (2020). Improving care for children affected by adverse childhood experiences. (Unpublished doctoral dissertation, East Carolina University College of Nursing). http://hdl.handle.net/10342/8660

López, V., Murphy, M., Lucke, C., Torres-Vallejos, J., Villalobos-Parada, B., Ascorra, P., ... & Bilbao, M. (2018). Peer Victimization and Mental Health Risk in Chilean Students. Journal of Child and Family Studies, 27(8), 2608-2621.

Metwally, A.M., El-Sonbaty, M.M., El Etreby, L.A., El-Din, E.M.S., Hamid, N.A., Hussien, H.A., Hassanin, A.M., & Monir, Z.M. (20//20). Impact of national Egyptian school feeding program on growth, development, and school achievement of school children. World Journal of Pediatrics, 16: 393-400. https://doi.org/10.1007/s12519-020-00342-8

Nasir, A., Zimmer, A., Taylor, D., & Santo, J. (2019). Psychosocial assessment of the family in the clinical setting. BMC psychology, 7(1), 3. doi:10.1186/s40359-018-0277-5

Roby, E., Piccolo, L. R., Gutierrez, J., Kesoglides, N. M., Raak, C. D., Mendelsohn, A. L., & Canfield, C. F. (2021). Father involvement in infancy predicts behavior and response to chronic stress in middle childhood in a low-income Latinx sample. Developmental Psychobiology. https://doi.org/10.1002/dev.22081

Suku, S., Soni, J., Martin, M. A., Mirza, M. P., Glassgow, A. E., Gerges, M., ... & Caskey, R. (2019). A multivariable analysis of childhood psychosocial behaviour and household functionality. Child: care, health and development, 45(4), 551-558.

Vargas, B., Leiva, L., Rojas-Andrade, R., & Scquicciarini, A.M. (2019). Effects of psychosocial adversity on school maladjustment: A follow-up of primary school students. Psicología Educativa, 25(2), 101-108. https://doi.org/10.5093/psed2019a10

Weigel, M. M., & Armijos, R. X. (2018). Household food insecurity and psychosocial dysfunction in Ecuadorian elementary schoolchildren. International Journal of Pediatrics. doi: 10.1155/2018/6067283

Menarche

Irannezhad, E. & Soltanizadeh, M. (2021). Comparison of mental health of female students with early, late, and normal menarche in Isfahan, Iran [published in Arabic].

Pediatric Hospitalizations

Doupnik, S. K., Henry, M. K., Bae, H., Litman, J., Turner, S., Scharko, A. M., & Feudtner, C. (2017). Mental Health Conditions and Symptoms in Pediatric Hospitalizations: A Single-Center Point Prevalence Study. Academic Pediatrics, 17(2), 184-190.

Restless Legs Syndrome

Stubbs, P.H. & Walters, A.S. (2020). Tools for the assessment of pediatric restless legs syndrome. Frontiers in Psychiatry, 11(356). 1-32. doi: 10.3389/fpsyt.2020.00356

Sleep Difficulties

Kim, E., Lee, R. & Cain, K. C. (2017). Cosleeping, sleep disturbances, children's behavioral problems, and parenting self-efficacy among Korean American families. Journal of Child & Adolescent Psychiatric Nursing, 30, 112-120. doi:10.1111/jcap.12182.

Lin, Y.-M., Kuo, S.-Y., Chang, Y.-K., Lin, P.-C., Lin, Y.-K., Lee, P.-H., Lin, P.-H., & Chen, S.-R. (2021). Effects of Parental Education on Screen Time, Sleep Disturbances, and Psychosocial Adaptation Among Asian Preschoolers: A Randomized Controlled Study. Journal of Pediatric Nursing-Nursing Care of Children & Families, 56, E27-E34. https://doi.org/10.1016/j.pedn.2020.07.003

Basch, M.C., Stromberg, S.E., Krietsch, K., Chardon, M.L., Reynolds, C.M., Acharya, R., & Janicke, D.M. (2019). Adolescent-reported sleep/wake patterns in the relationships between inhibitory control and internalizing and externalizing problems. Journal of Developmental and Behavioral Pediatrics, 40(9). 679-685.

Thalassemia

Elzaree, F. A., Shehata, M. A., El Wakeel, M. A., El-Alameey, I. R., AbuShady, M. M., & Helal, S. I. (2018). Adaptive Functioning and Psychosocial Problems in Children with Beta Thalassemia Major. Open Access Macedonian Journal of Medical Sciences, 6(12), 2337–2341. doi:10.3889/oamjms.2018.367

Raman, V., Prakash, A., & D'Souza, F. (2019). Psychosocial Issues in Children With Thalassemia: From Identification to a Model for Management in a Developing Country. Journal of Pediatric Hematology/Oncology, 41(3), 218-221.

Ghorbanpoor, M., Mirzaie, M., Mirhaghjou, S.N., & Roshan, Z.A. (2020). The relationship between psychosocial status and adherence to treatment regimen in adolescents with thalassemia. Journal of Holistic Nursing and Midwifery, 30(2). 78-85. DOI: 10.32598/jhnm.30.2.78

School climate

Sajady, M., Gower, A. L., McCullough, M., & Jordan, C. (2020). More than a View: School Landscape Features Are Associated with Improved Student Adjustment. Journal of Developmental and Behavioral Pediatrics, 41(6), 436-442. https://doi.org/10.1097/dbp.0000000000000809

Substance abuse

Anyanwu, O. U., Ibekwe, R. C., & Ojinnaka, N. C. (2017). Psychosocial dysfunction among adolescents who abuse substances in secondary schools in Abakaliki, Nigeria. Nigerian Journal Of Clinical Practice, 20(6), 665-669. doi:10.4103/1119-3077.208953.

Shahzad, S., Kliewer, W., Ali, M., & Begum, N. (2020). Different risk factors are associated with alcohol use versus problematic use in male Pakistani adolescents. International Journal of Psychology, 55(4), 585-589. https://doi.org/10.1002/ijop.12617

Weight Management

Gowey, M., Redden, D., Lim, C., Janicke, D., & Dutton, G. (2019). Executive function phenotypes in pediatric obesity. Pediatric Obesity. https://doi.org/10.1111/ijpo.12655

Ratcliff, M. B., Catlin, P. A., Peugh, J. L., Siegel, R. M., Kirk, S., & Tamm, L. (2017). Psychosocial Screening Among Youth Seeking Weight Management Treatment. Clinical Pediatrics, 0009922817715936.

Thomaseo Burton, E. T., Jones, T. L., Smith, W. A., & Han, J. C. (2019) Psychosocial Functioning After One Year of Interdisciplinary Pediatric Weight Management, Behavioral Medicine, doi: 10.1080/08964289.2019.1570073

Wylie-Rosett, J., Groisman-Perelstein, A. E., Diamantis, P. M., Jimenez, C. C., Shankar, V., Conlon, B. A., ... & Matthan, N. R. (2018). Embedding weight management into safety-net pediatric primary care: randomized controlled trial. International Journal of Behavioral Nutrition and Physical Activity, 15(1), 12.

Other Chronic Conditions

Hastuti et al. (2020). The effect of short stature on children's cognitive abilities and psychosocial condition. Journal of Maternal and Child Health, 5(3). 297-303.

Im, Y., & Kim, D. H. (2021). Family Management Style and Psychosocial Health of Children with Chronic Conditions. Journal of Child and Family Studies, 30(2), 483-492. https://doi.org/10.1007/s10826-020-01870-7

Rukabyarwema, J.P., McCall, N., Ngambe, T., Kanyembari, X.B., & Needlman, R. (2019). Behavioral problems in physically ill children in Rwanda. Journal of Developmental and Behavioral Pediatrics, 40(8). 642-650. doi: 10.1097/DBP.0000000000000698

COVID-19

Bate, J., Pham, P. T., & Borelli, J. L. (2021). Be My Safe Haven: Parent-Child Relationships and Emotional Health During COVID-19 [Article]. Journal of Pediatric Psychology, 46(6), 624-634. https://doi.org/10.1093/jpepsy/jsab046

Bourne, T. Y., Kumar, S., & Coutu, F. H. (2021). COVID-19 Effects on Education and Mental Health within Adolescents. In: Am Acad Pediatrics.

Hussong, A. M., Midgette, A. J., Thomas, T. E., Coffman, J. L., & Cho, S. (2021). Coping and Mental Health in Early Adolescence during COVID-19. Research on Child and Adolescent Psychopathology. https://doi.org/10.1007/s10802-021-00821-0

Lund, Emily & Gabrielli, Joy. (2021). The Role of Pediatric Psychologists in Mitigating Disability-Specific Barriers Among Youth During the COVID-19 Pandemic. Clinical Practice in Pediatric Psychology, 9, 12-23. https://doi.org/10.1037/cpp0000387

Mekori-Domachevsky, E., Matalon, N., Mayer, Y., Shiffman, N., Lurie, I., Gothelf, D., & Dekel, I. Internalizing symptoms impede adolescents' ability to transition from in-person to online mental health services during the 2019 coronavirus disease pandemic. Journal of Telemedicine and Telecare, Article 1357633x211021293. https://doi.org/10.1177/1357633x211021293

Oliva, S., Russo, G., Gili, R., Russo, L., Di Mauro, A., Spagnoli, A., Alunni Fegatelli, D., Romani, M., Costa, A., & Veraldi, S. (2021). Risks and Protective Factors Associated With Mental Health Symptoms During COVID-19 Home Confinement in Italian Children and Adolescents: The# Understandingkids Study. Frontiers in Pediatrics, 9, 537.

Ozturk, F. O., & Ayaz-Alkaya, S. (2021). Internet addiction and psychosocial problems among adolescents during the COVID-19 pandemic: A cross-sectional study [Article]. Archives of Psychiatric Nursing, 35(5), 595-601. https://doi.org/10.1016/j.apnu.2021.08.007

Riffat, O., Humayun Iqbal, K., Muhammad Khalid, M., Najaf, M., & Fatima, T. (2021). Psychosocial impact of the COVID-19 pandemic on doctors' children: are we heading towards a mental health pandemic? Paediatrica Indonesiana, 61(1), 46-52. https://doi.org/10.14238/pi61.1.2021.46-52

Shah, S., Kaul, A., Shah, R., & Maddipoti, S. (2021). Impact of Coronavirus Disease 2019 Pandemic and Lockdown on Mental Health Symptoms in Children. Indian Pediatrics, 58(1), 75-76.

II. Literature Reviews

Acra, P. E., & Pérez, M. T. (2017). The evaluation of adolescent chest pain: a screening ECG or PSC-17? Current Opinion In Pediatrics, 29(4), 414-419. doi:10.1097/MOP.0000000000000503.

Becker-Haimes, E.M., Tabachnick, A.R., Last, B.S., Stewart, R.E., Hasan-Granier, A., & Beidas, R.S. (2019). Evidence base update for brief, free, and accessible youth mental health measures. Journal of Clinical Child & Adolescent Psychology, 49. 1-17. https://doi.org/10.1080/15374416.2019.1689824

Black, J. J., & Rofey, D. L. (2018). An overview of common psychiatric problems among adolescent and young adult females: focus on mood and anxiety. Best Practice & Research Clinical Obstetrics & Gynaecology, 48, 165-173.

Brino, K.A.S. (2020). Pediatric mental health and the power of primary care: Practical approaches and validating challenges. Journal of Pediatric Health Care, 34(2). 12-20.

Bultas, M. W., McMillin, S. E., Broom, M. A., & Zand, D. H. (2017). Brief, Rapid Response, Parenting Interventions Within Primary Care Settings. J Behav Health Serv Res, 44(4), 695-699. https://doi.org/10.1007/s11414-015-9479-2

Burt, J., Clarke, B., & LaBrot, Z. (2018). Screening and Assessment of Disruptive, Impulse-Control, and Conduct Disorders. In Handbook of Pediatric Psychological Screening and Assessment in Primary Care (pp. 395-423). Routledge.

Caballero, T. M., DeCamp, L. R., Platt, R. E., Shah, H., Johnson, S. B., Sibinga, E. M., & Polk, S. (2017). Addressing the Mental Health Needs of Latino Children in Immigrant Families. Clinical pediatrics, 56(7), 648-658.

Cianchetti, C. (2020). Early detection of behavioral and emotional problems in school-aged children and adolescents: The parent questionnaires. Clinical Practice & Epidemiology in Mental Health, 16, 7-16. doi: 10.2174/1745017902016010007

D'Angelo, E., Tunick, R., Sinclair-McBride, K., & Morelli, N. (2018). Global Screening for Behavioral Health Disorders. In Handbook of Pediatric Psychological Screening and Assessment in Primary Care (pp. 134-162). Routledge.

Forman-Hoffman, V. L., & Viswanathan, M. (2018). Screening for Depression in Pediatric Primary Care. Current psychiatry reports, 20(8), 62.

Hiott, D. B., Phillips, S., & Amella, E. (2018). Adolescent risk screening instruments for primary care: An integrative review utilizing the Donabedian Framework. Comprehensive Child & Adolescent Nursing, 41(4), 255-275. https://doi.org/10.1080/24694193.2017.1330372

Jeffrey, J., Klomhaus, A., Enenbach, M., Lester, P., & Krishna, R. (2020). Self-Report Rating Scales to Guide Measurement-Based Care in Child and Adolescent Psychiatry. Child and Adolescent Psychiatric Clinics of North America, 29(4), 601-629. https://doi.org/10.1016/j.chc.2020.06.002

Jellinek, M. (2021). The Pediatric Symptom Checklist: A Bridge to Child and Adolescent Psychiatry from Pediatrics. Journal of the American Academy of Child & Adolescent Psychiatry. 2005

Jellinek, M., Murphy, J. M. (2020). Screening for psychosocial functioning as the eighth vital sign. JAMA Pediatrics. Published online July 20, 2020. doi:10.1001/jamapediatrics.2020.2005

Joshi, H., & Apple, R. (2019). Screening Adolescents for ADHD, Oppositional Defiant Disorder, and Conduct Disorder in Primary Care. In V. Morelli (Ed.), Adolescent Health Screening: an Update in the Age of Big Data (pp. 107-113). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-323-66130-0.00009-0

Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries. Autism Research, 12(2), 176-199.

Marti, F. A., Pourat, N., Lee, C., & Zima, B. T. (2021). Overview of Ten Child Mental Health Clinical Outcome Measures: Testing of Psychometric Properties with Diverse Client Populations in the US. Administration and Policy in Mental Health and Mental Health Services Research. https://doi.org/10.1007/s10488-021-01157-z

Pourat, N., Zima, B., Marti, A., & Lee, C. (2017). California child mental health performance outcomes system: Recommendation report. UCLA Center for Health Policy Research.

Simha, S., & Brown, A. C. (2021). Preventive Care in Children and Adolescents. Primary Care: Clinics in Office Practice, 48(1), 99-116.

Thabrew, H., McDowell, H., Given, K., & Murrell, K. (2017). Systematic Review of Screening Instruments for Psychosocial Problems in Children and Adolescents With Long-Term Physical Conditions. Global Pediatric Health, 4, 2333794X17690314.

Trafalis, S., Giannini, C., Joves, J., Portera, S., Toyama, H., Mehta, A., Basile, K., & Friedberg, R. D. (2021). A pediatrician-friendly review of three common behavioral health screeners in pediatric practice: Findings and recommendations. Pediatric Investigation, 5(1), 58-64. https://doi.org/10.1002/ped4.12246

Zima, B. T., Marti, F. A., Lee, C. E., & Pourat, N. (2019). Selection of a Child Clinical Outcome Measure for Statewide Use in Publicly Funded Outpatient Mental Health Programs. Journal of Psychiatric Services, 70(5), 381-388. doi: 10.1176/appi.ps.201800424

2. References for all papers cited in 2021 (and 2017) NQF submissions

1. Achenbach TM. *Manual for the Child Behavior Checklist and 1991 Profile*. University of Vermont, Department of Psychiatry; 1991.

2. Acra PE, Perez MT. The evaluation of adolescent chest pain: a screening ECG or PSC-17? Review. *Current opinion in pediatrics*. Aug 2017;29(4):414-419. doi:10.1097/MOP.0000000000000503

3. Afzal E, Khan WI, Iqbal MK, Aftab K, Ahmad T. Prevalence of psychosocial problems in children with and without epilepsy: A comparative cross-section study. *The Professional Medical Journal*. 2021;28(04):557-563. doi:https://doi.org/10.29309/TPMJ/2021.28.04.5800

4. American Academy of Pediatrics. *Addressing Mental Health Concerns in Primary Care: A Clinician's Toolkit*. 2015.

5. American Academy of Child and Adolescent Psychiatry Committee on Health Care Access and Economics Task Force on Mental Health. Improving mental health services in primary care: Reducing administrative and financial barriers to access and collaboration *Pediatrics* 2009;123 (4):1248-51.

6. American Academy of Pediatrics. AAP Publications Reaffirmed or Retired. *Pediatrics*. 2021;147(6)doi:https://doi.org/10.1542/peds.2021-051488

7. Anderson RL, Lyons JS, Giles DM, Price JA, Estle G. Reliability of the child and adolescent needs and strengths-mental health (CANS-MH) scale. *Journal of Child and Family Studies*. 2003;12(3):279-289.

8. Anker MG, Duncan BL, Sparks JA. Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of consulting and clinical psychology*. 2009;77(4):693.

9. Anyanwu OU, Ibekwe RC, Ojinnaka NC. Psychosocial dysfunction among adolescents who abuse substances in secondary schools in Abakaliki, Nigeria. *Niger J Clin Pract*. 06 2017;20(6):665-669. doi:10.4103/1119-3077.208953

10. Arauz-Boudreau A, Riobueno-Naylor A, Haile H, et al. How an Electronic Medical Record System Facilitates and Demonstrates Effective Psychosocial Screening in Pediatric Primary Care. *Clin Pediatr (Phila)*. 02 2020;59(2):154-162. doi:10.1177/0009922819892038

11. Ardıç E, Ünsal G, Bayram S. Validity and reliability of the Pictorial Pediatric Symptom Checklist. *Journal of Psychiatric Nursing*. 2020;11(1):20-27.

12. Basch MC, Stromberg SE, Krietsch K, et al. Adolescent-Reported Sleep/Wake Patterns in the Relationships Between Inhibitory Control and Internalizing and Externalizing Problems. *Journal of developmental and behavioral pediatrics : JDBP*. 12 2019;40(9):679-685. doi:10.1097/dbp.0000000000000707

13. Bate J, Pham PT, Borelli JL. Be My Safe Haven: Parent-Child Relationships and Emotional Health During COVID-19. Article. *Journal of Pediatric Psychology*. Jul 2021;46(6):624-634. doi:10.1093/jpepsy/jsab046

14. Baugh EJ, Ballard SM, Carter L, Tyndall L, Nolan M, Littlewood K. The County-Wide Dissemination of Triple P Parenting Program: An Implementation Study. *Prog Community Health Partnersh*. 2019;13(1):73-81. doi:10.1353/cpr.2019.0010

15. Becker-Haimes EM, Tabachnick AR, Last BS, Stewart RE, Hasan-Granier A, Beidas RS. Evidence Base Update for Brief, Free, and Accessible Youth Mental Health Measures. Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*. Jan-Feb 2020;49(1):1-17. doi:10.1080/15374416.2019.1689824

16. Berger-Jenkins E, McCord M, Gallagher T, Olfson M. Effect of routine mental health screening in a low-resource pediatric primary care population. *Clinical pediatrics*. 2012;51(4):359-365.

17. Berger-Jenkins E, Monk C, D'Onfro K, et al. Screening for Both Child Behavior and Social Determinants of Health in Pediatric Primary Care. Research Support, Non-U.S. Gov't. *Journal of developmental and behavioral pediatrics : JDBP*. Jul/Aug 2019;40(6):415-424. doi:10.1097/DBP.0000000000000676

18. Bergmann P, Lucke C, Nguyen T, Jellinek M, Murphy JM. Identification and utility of a short form of the Pediatric Symptom Checklist-Youth self-report (PSC-17-Y). *European Journal of Psychological Assessment*. 2020;36(1):56-64. doi:10.1027/1015-5759/a000486

19. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*. 2011;62(12):1423-1429.

20. Black JJ, Rofey DL. An overview of common psychiatric problems among adolescent and young adult females: Focus on mood and anxiety. *Best Pract Res Clin Obstet Gynaecol*. Apr 2018;48:165-173. doi:10.1016/j.bpobgyn.2017.08.007

21. Blucker RT, Jackson D, Gillaspy JA, Hale J, Wolraich M, Gillaspy SR. Pediatric Behavioral Health Screening in Primary Care A Preliminary Analysis of the Pediatric Symptom Checklist-17 With Functional Impairment Items. *Clinical pediatrics*. 2014;53(5):449-455.

22. Boothroyd RA, Armstrong M. An examination of the psychometric properties of the Pediatric Symptom Checklist with children enrolled in Medicaid. *Journal of Emotional and Behavioral Disorders*. 2010;18(2):113-126.

23. Bourne TY, Kumar S, Coutu FH. COVID-19 Effects on Education and Mental Health within Adolescents. *Pediatrics*. 2021;147(3 MeetingAbstract):213-213. doi:10.1542/peds.147.3_MeetingAbstract.213

24. Brino KAS. Pediatric Mental Health and the Power of Primary Care: Practical Approaches and Validating Challenges. *J Pediatr Health Care*. 2020 Mar - Apr 2020;34(2):e12-e20. doi:10.1016/j.pedhc.2019.09.013

25. Brodey BB, Cuffel B, McCulloch J, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *The American journal of managed care*. 2005;11(12):774-780.

26. Buchbinder D, Patel SK, Casillas JN, et al. Parent proxy assessment of sibling quality of life following pediatric hematopoietic cell transplantation. *Health Qual Life Outcomes*. Oct 29 2019;17(1):162. doi:10.1186/s12955-019-1231-9

27. Burdzovic Andreas J, O'Farrell TJ. Psychosocial problems in children of women entering substance use disorder treatment: A longitudinal study. Research Support, U.S. Gov't, Non-P.H.S. Research Support, N.I.H., Extramural. *Addictive behaviors*. Feb 2017;65:193-197. doi:10.1016/j.addbeh.2016.10.015

28. Burke RE, Hoffman ND, Guy L, Bailey J, Silver EJ. Screening, Monitoring, and Referral to Treatment for Young Adolescents at an Urban School-Based Health Center. *Journal of School Health*. 2021;91(11)doi:https://doi.org/10.1111/josh.13089

29. Burt J, Clarke B, LaBrot Z. Screening and Assessment of Disruptive, Impulse-Control, and Conduct Disorders. *Handbook of Pediatric Psychological Screening and Assessment in Primary Care*. Routledge; 2018:395-423.

30. Caballero TM, DeCamp LR, Platt RE, et al. Addressing the Mental Health Needs of Latino Children in Immigrant Families. *Clin Pediatr (Phila)*. Jun 2017;56(7):648-658. doi:10.1177/0009922816679509

31. Chaffin M, Campbell C, Whitworth DN, et al. Accuracy of a Pediatric Behavioral Health Screener to Detect Untreated Behavioral Health Problems in Primary Care Settings. *Clin Pediatr (Phila)*. May 2017;56(5):427-434. doi:10.1177/0009922816678412

32. Chaurasiya A, Ranjan JK, Asthana HS. Hindi adaptation of Pediatric Symptom Checklist-Youth Version (PSC-Y). *Journal of Indian Association for Child & Adolescent Mental Health*. 2019;15(1):69-84.

33. Chellan D, Muktesh G, Vaiphei K, et al. Effect of gluten-free diet and compliance on quality of life in pediatric celiac disease patients. *JGH Open*. Oct 2019;3(5):388-393. doi:10.1002/jgh3.12172

34. Choi KR, Stewart T, Fein E, et al. The Impact of Attachment-Disrupting Adverse Childhood Experiences on Child Behavioral Health. *J Pediatr*. 06 2020;221:224-229. doi:10.1016/j.jpeds.2020.03.006

35. Choudhary S, Niranjan N, Khichar S, Berwal PK, Barath AS. Behavioral Problems and Intelligence Quotient Changes in Pediatric Epilepsy: A Case-Control Study. *J Neurosci Rural Pract*. Oct-Dec 2017;8(4):617-621. doi:10.4103/jnrp.jnrp_57_17

36. Cianchetti C. Early Detection of Behavioral and Emotional Problems in School-Aged Children and Adolescents: The Parent Questionnaires. *Clin Pract Epidemiol Ment Health*. 2020;16:7-16. doi:10.2174/1745017902016010007

37. Committee on Psychosocial Aspects of Child Family Health, Task Force on Mental Health. The Future of Pediatrics: Mental Health Competencies for Pediatric Primary Care. *Pediatrics*. 2009;124(1):410-421. doi:10.1542/peds.2009-1061

38. Costello EJ, Costello AJ, Edelbrock C, et al. Psychiatric disorders in pediatric primary care: Prevalence and risk factors. *Archives of General Psychiatry*. 1988;45(12):1107-1116.

39. Craig R, Naganathan S, Ballance C, Schairer J, Fleur RS. Improving Rates and Referrals of Developmental and Mental Health Screening at Well Child Visits. Am Acad Pediatrics; 2018.

40. D'Angelo E, Tunick R, Sinclair-McBride K, Morelli N. Global Screening for Behavioral Health Disorders. *Handbook of Pediatric Psychological Screening and Assessment in Primary Care*. 1 ed. Routledge; 2018.

41. DiStefano C, Barth SG, Greer F. Assessing Equivalency of PSC-17 Ratings: Does It Matter if Mixed or Grouped Item Format Is Used? *Journal of Psychoeducational Assessment*. 10/01/ 2019;37(7):920-924.

42. DiStefano C, Liu J, Burgess Y. Investigating the Structure of the Pediatric Symptoms Checklist in the Preschool Setting. *Journal of Psychoeducational Assessment*. 2017;35(5):494-505. doi:10.1177/0734282916647648

43. Doupnik SK, Henry MK, Bae H, et al. Mental Health Conditions and Symptoms in Pediatric Hospitalizations: A Single-Center Point Prevalence Study. *Acad Pediatr*. 03 2017;17(2):184-190. doi:10.1016/j.acap.2016.08.009

44. Dowrick C, Leydon GM, McBride A, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *Bmj*. 2009;338:b663.

45. El-Sonbaty MM, Fathy A, Aljohani A. Assessment of Behavioural Disorders in Children with Functional Constipation. *Open Access Maced J Med Sci*. Dec 15 2019;7(23):4019-4022. doi:10.3889/oamjms.2019.677

46. Elzaree FA, Shehata MA, El Wakeel MA, El-Alameey IR, AbuShady MM, Helal SI. Adaptive Functioning and Psychosocial Problems in Children with Beta Thalassemia Major. *Open Access Maced J Med Sci*. Dec 20 2018;6(12):2337-2341. doi:10.3889/oamjms.2018.367

47. Fallucco EM, Wysocki T, James L, Kozikowski C, Williams A, Gleason MM. The Brief Early Childhood Screening Assessment: Preliminary Validity in Pediatric Primary Care. *Journal of developmental and behavioral pediatrics : JDBP*. 2017 Feb/Mar 2017;38(2):89-98. doi:10.1097/dbp.0000000000000384

48. Fonseca A, Kral MC, Lally MD, Boan AD, Macias MM. Bullying and ostracism in youth with and without ADHD: Implications for risk and resilience. *Children's Health Care*. 2019;48(2):164-175. doi:10.1080/02739615.2018.1514607

49. Force UPST. Screening and treatment for major depressive disorder in children and adolescents: US Preventive Services Task Force Recommendation Statement. *Pediatrics*. 2009;123(4):1223-1228.

50. Forman-Hoffman V, McClure E, McKeeman J, et al. Screening for Major Depressive Disorder Among Children and Adolescents: A Systematic Review for the U.S. Preventive Services Task Force. *Evidence Synthesis No 116 AHRQ Publication No 13-05192-EF-1 Rockville, MD: Agency for Healthcare Research and Quality*. 2016;

51. Forman-Hoffman V, McClure E, McKeeman J, et al. Screening for Major Depressive Disorder in Children and Adolescents: A Systematic Review for the US Preventive Services Task ForceScreening for MDD in Children and Adolescents. *Annals of internal medicine*. 2016;164(5):342-349.

52. Forman-Hoffman VL, Viswanathan M. Screening for Depression in Pediatric Primary Care. *Curr Psychiatry Rep*. 07 23 2018;20(8):62. doi:10.1007/s11920-018-0926-7

53. Fortney J, Sladek R, Unützer J, et al. *Fixing Behavioral Health Care in America: A National Call for Measurement-Based Care in the Delivery of Behavioral Health Service*. 2015. *The Kennedy Forum*.

54. Fortney J, Sladek R, Unützer J, et al. *A Core Set of Outcome Measures For Behavioral Health Across Service Settings Supplement to Fixing Behavioral Health Care in America: A National Call for Measurement-Based Care in the Delivery of Behavioral Health Services*. 2015.

55. Foy JM, Green CM, Earls MF, CHILD COPAO, FAMILY HEALTH MHLWG. Mental Health Competencies for Pediatric Practice. *Pediatrics*. 2019;144(5):e20192757. doi:10.1542/peds.2019-2757

56. Foy JM, Kelleher KJ, Laraque D, Health AAoPTFoM. Enhancing pediatric mental health care: strategies for preparing a primary care practice. *Pediatrics*. 2010;125(Supplement 3):S87-S108.

57. Gaete J. Validation of the Paediatric Symptom Checklist to Identify and Assess Psychosocial Problems among Early Adolescents in Chile, 2017. Data. *UK Data Archive*. 2021;doi:http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.5255/UKDA-SN-854843

58. Gall G, Pagano ME, Desmond MS, Perrin JM, Murphy JM. Utility of Psychosocial Screening at a School-Based Health Center. *Journal of School Health*. 2000;70(7):292-298.

59. Gardner W, Bevans K, Kelleher KJ. The Potential for Improving the Population Health Effectiveness of Screening: A Simulation Study. *Pediatrics*. 2021;148(Supplement 1):s3-s10. doi:https://doi.org/10.1542/peds.2021-050693C

60. Gardner W, Lucas A, Kolko DJ, Campo JV. Comparison of the PSC-17 and alternative mental health screens in an at-risk primary care sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2007;46(5):611-618.

61. Gardner W, Murphy M, Childs G, et al. The PSC-17: A brief pediatric symptom checklist with psychosocial problem subscales. *Ambulatory Child Health*. 1999;5:225-236.

62. Ghorbanpoor M, Mirzaie M, Mirhaghjou SN, Atrkar Roshan Z. The Relationship Between Psychosocial Status and Adherence to Treatment Regimen in Adolescents With Thalassemia. Research. *Journal of Holistic Nursing and Midwifery*. 2020;30(2):78-85. doi:10.32598/jhnm.30.2.78

63. Goldstein LA, Gibbons MBC, Thompson SM, et al. Outcome assessment via handheld computer in community mental health: Consumer satisfaction and reliability. *The journal of behavioral health services & research*. 2011;38(3):414-423.

64. Goodman R. The Strengths and Difficulties Questionnaire: A Research Note. *J Child Psychol Psychiatry*. 1997;38:581-586.

65. Gowey M, Redden D, Lim C, Janicke D, Dutton G. Executive function phenotypes in pediatric obesity. *Pediatr Obes*. 09 2020;15(9):e12655. doi:10.1111/ijpo.12655

66. Green M. *Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents*. ERIC; 1994.

67. Groves AP, Gettinger K, Druley TE, et al. Special Therapy and Psychosocial Needs Identified in a Multidisciplinary Cancer Predisposition Syndrome Clinic. *J Pediatr Hematol Oncol*. 03 2019;41(2):133-136. doi:10.1097/mph.0000000000001251

68. Guzmán J, Kessler RC, Squicciarini AM, et al. Evidence for the effectiveness of a national school-based mental health program in Chile. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2015;54(10):799-807. e1.

69. Guzman MP, Jellinek M, George M, et al. Mental health matters in elementary school: First-grade screening predicts fourth grade achievement test scores. *European child & adolescent psychiatry*. 2011;20(8):401-411.

70. Hacker KA, Arsenault L, Franco I, et al. Referral and follow-up after mental health screening in commercially insured adolescents. *Journal of Adolescent Health*. 2014;55(1):17-23.

71. Hacker KA, Arsenault LN, Williams S, DiGirolamo AM. Mental and behavioral health screening at preventive visits: opportunities for follow-up of patients who are nonadherent with the next preventive visit. *The Journal of Pediatrics*. 2011;158(4):666-671. e2.

72. Hacker KA, Goldstein J, Link D, et al. Pediatric provider processes for behavioral health screening, decision making, and referral in sites with colocated mental health services. *Journal of Developmental & Behavioral Pediatrics*. 2013;34(9):680-687.

73. Hacker KA, Myagmarjav E, Harris V, Suglia SF, Weidner D, Link D. Mental health screening in pediatric practice: factors related to positive screens and the contribution of parental/personal concern. *Pediatrics*. 2006;118(5):1896-1906.

74. Hacker KA, Penfold R, Arsenault L, Zhang F, Murphy M, Wissow L. Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. *Pediatrics*. 2014;133(1):46-54.

75. Hacker KA, Penfold R, Arsenault LN, Zhang F, Soumerai SB, Wissow LS. The Impact of the Massachusetts Behavioral Health Child Screening Policy on Service Utilization. *Psychiatric Services*. 2016;68(1):25-32.

76. Hacker KA, Penfold RB, Arsenault LN, Zhang F, Murphy M, Wissow LS. Behavioral health services following implementation of screening in Massachusetts Medicaid children. *Pediatrics*. 2014:peds. 2014-0453.

77. Hacker KA, Williams S, Myagmarjav E, Cabral H, Murphy M. Persistence and change in pediatric symptom checklist scores over 10 to 18 months. *Academic Pediatrics*. 2009;9(4):270-277.

78. Hansel T, Rohrer G, Osofsky J, Osofsky H, Arthur E, Barker C. Integration of Mental and Behavioral Health in Pediatric Health Care Clinics. *J Public Health Manag Pract*. 2017 Nov/Dec 2017;23 Suppl 6 Suppl, Gulf Region Health Outreach Program:S19-S24. doi:10.1097/phh.0000000000000649

79. Harmon SC, Lambert MJ, Smart DM, et al. Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy research*. 2007;17(4):379-392.

80. Hastuti H, Nisa Z, Umma H, Putri A, Mashuri Y. The Effect of Short Stature on Children's Cognitive Abilities and Psychosocial Condition. *Journal of Maternal and Child Health*. 01/01 2020;5:297-303. doi:10.26911/thejmch.2020.05.03.08

81. Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy research*. 2004;14(3):308-327.

82. Hawkins J. Improving Care for Children Affected by Adverse Childhood Experiences. DNP Scholarly Project: East Carolina University; 2020.

83. Hayutin LG, Reed-Knight B, Blount RL, Lewis J, McCormick ML. Increasing parent–pediatrician communication about children's psychosocial problems. *Journal of Pediatric Psychology*. 2009;34(10):1155-1164.

84. Higuchi T, Ishizaki Y, Uenishi H, et al. Reliability and Validity of the Japanese Version of the Youth-rating Pediatric Symptom Checklist Shortened Version. 2021;

85. Hiott DB, Phillips S, Amella E. Adolescent Risk Screening Instruments for Primary Care: An Integrative Review Utilizing the Donabedian Framework. *Comprehensive Child & Adolescent Nursing*. 2018;41(4):255-275. doi:10.1080/24694193.2017.1330372

86. Hogan M. The President's New Freedom Commission: recommendations to transform mental health care in America. . *Psychiatric Services* 2003;54:1467-1474.

87. Holcomb JM, Boudreau AA, Riobueno-Naylor A, et al. Beyond Initial Screening: One-Year Follow-up of Adolescents with Internalizing Problems on the Pediatric Symptom Checklist. *Journal of Developmental & Behavioral Pediatrics*. 2021;

88. Honigfeld L, Macary SJ, Grasso DJ. A Clinical Care Algorithmic Toolkit for Promoting Screening and Next-Level Assessment of Pediatric Depression and Anxiety in Primary Care. Article. *Journal of Pediatric Health Care*. 2017;31(3):e15-e23. doi:10.1016/j.pedhc.2017.01.008

89. Huang KY, Bornheimer LA, Dankyi E, de-Graft Aikins A. Parental Wellbeing, Parenting and Child Development in Ghanaian Families with Young Children. *Child Psychiatry Hum Dev*. 10 2018;49(5):833-841. doi:10.1007/s10578-018-0799-3

90. Hussong AM, Midgette AJ, Thomas TE, Coffman JL, Cho S. Coping and Mental Health in Early Adolescence during COVID-19. *Research on Child and Adolescent Psychopathology*. 2021;doi:10.1007/s10802-021-00821-0

91. Im Y, Kim DH. Family Management Style and Psychosocial Health of Children with Chronic Conditions. *Journal of child and family studies*. Feb 2021;30(2):483-492. doi:10.1007/s10826-020-01870-7

92. Irannezhad E, Soltanizadeh M. Comparison of Mental Health of Female Students with Early, Late, and Normal Menarche in Isfahan, Iran. 03/01/21 2021;27(4):253-260. doi:10.29252/ajcm.27.4.253

93. Jacobson JH, Pullmann MD, Parker EM, Kerns SEU. Measurement Based Care in Child Welfare-Involved Children and Youth: Reliability and Validity of the PSC-17. *Child Psychiatry Hum Dev*. 04 2019;50(2):332-345. doi:10.1007/s10578-018-0845-1

94. Jeffrey J, Klomhaus A, Enenbach M, Lester P, Krishna R. Self-Report Rating Scales to Guide Measurement-Based Care in Child and Adolescent Psychiatry. *Child and adolescent psychiatric clinics of North America*. Oct 2020;29(4):601-629. doi:10.1016/j.chc.2020.06.002

95. Jellinek M, Bergmann P, Holcomb JM, et al. Recognizing Adolescent Depression with Parent-and Youth-Report Screens in Pediatric Primary Care. *The Journal of Pediatrics*. 2021;233:220-226. e1.

96. Jellinek MS, Murphy JM, Burns BJ. Brief psychosocial screening in outpatient pediatric practice. *The Journal of Pediatrics*. 1986;109(2):371-378.

97. Jellinek MS, Murphy JM, Little M, Pagano ME, Comer DM, Kelleher KJ. Use of the Pediatric Symptom Checklist to screen for psychosocial problems in pediatric primary care: a national feasibility study. *Archives of pediatrics & adolescent medicine*. 1999;153(3):254-260.

98. Jellinek MS, Murphy JM, Robinson J, Feins A, Lamb S, Fenton T. Pediatric Symptom Checklist: screening school-age children for psychosocial dysfunction. *The Journal of Pediatrics*. 1988;112(2):201-209.

99. Jonovich SJ, Alpert-Gillis LJ. Impact of pediatric mental health screening on clinical discussion and referral for services. *Clinical pediatrics*. 2013:0009922813511146.

100. Joshi D, Tiwari MK, Kannan V, Dalal SS, Mathai SS. Emotional and behavioral disturbances in school going HIV positive children attending HIV clinic. *Med J Armed Forces India*. Jan 2017;73(1):18-22. doi:10.1016/j.mjafi.2016.12.002

101. Joshi H, Apple R. Screening Adolescents for ADHD, Oppositional Defiant Disorder, and Conduct Disorder in Primary Care. In: Morelli V, ed. *Adolescent Health Screening: an Update in the Age of Big Data*. Elsevier; 2019:107-113.

102. Kamin HS, McCarthy AE, Abel MR, Jellinek MS, Baer L, Murphy JM. Using a brief parent-report measure to track outcomes for children and teens with internalizing disorders. *Child Psychiatry & Human Development*. 2015;46(6):851-862.

103. Katzelnick DJ, Duffy FF, Chung H, Regier DA, Rae DS, Trivedi MH. Depression outcomes in psychiatric clinical practice: using a self-rated measure of depression severity. *Psychiatric Services*. 2011;62(8):929-935.

104. Kefale D, Boka A, Mengstu Z, Belayneh Z, Zeleke S. Emotional and behavioral problems and associated factors among children and adolescents on highly active anti-retroviral therapy in public hospitals of West Gojjam zone, Amhara regional state of Ethiopia, 2018: a cross-sectional study. *BMC Pediatr*. 05 03 2019;19(1):141. doi:10.1186/s12887-019-1453-3

105. Kellam SG, Wang W, Mackenzie AC, et al. The impact of the Good Behavior Game, a universal classroom-based preventive intervention in first and second grades, on high-risk sexual behaviors and drug abuse and dependence disorders into young adulthood. *Prevention Science*. 2014;15(1):6-18.

106. Kelleher KJ, Childs GE, Wasserman RC, McInerny TK, Nutting PA, Gardner WP. Insurance status and recognition of psychosocial problems: a report from the Pediatric Research in Office Settings and the Ambulatory Sentinel Practice Networks. *Archives of pediatrics & adolescent medicine*. 1997;151(11):1109-1115.

107. Kelleher KJ, McInerny TK, Gardner WP, Childs GE, Wasserman RC. Increasing identification of psychosocial problems: 1979–1996. *Pediatrics*. 2000;105(6):1313-1321.

108. Kieling C, Baker-Henningham H, Belfer M, et al. Child and adolescent mental health worldwide: evidence for action. *The Lancet*. 2011;378(9801):1515-1525.

109. Kim E, Lee R, Cain KC. Cosleeping, sleep disturbances, children's behavioral problems, and parenting self-efficacy among Korean American families. *J Child Adolesc Psychiatr Nurs*. May 2017;30(2):112-120. doi:10.1111/jcap.12182

110. King RB, Neilson E, King TM. Challenges and Opportunities to Bolster the Effectiveness of Childhood Screening. *Pediatrics*. 2021;148(Supplement 1):s51-s55. doi:https://doi.org/10.1542/peds.2021-050693K

111. Klages KL, Berlin KS, Silverman AH, et al. Empirically Derived Patterns of Pain, Stooling, and Incontinence and Their Relations to Health-Related Quality of Life Among Youth With Chronic Constipation. *Journal of Pediatric Psychology*. 2017;42(3)doi:https://doi.org/10.1093/jpepsy/jsw068

112. Knaup C, Koesters M, Schoefer D, Becker T, Puschner B. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *The British Journal of Psychiatry*. 2009;195(1):15-22.

113. Kolko DJ, Campo J, Kilbourne AM, Hart J, Sakolsky D, Wisniewski S. Collaborative care outcomes for pediatric behavioral health problems: a cluster randomized trial. *Pediatrics*. 2014:peds. 2013-2516.

114. Kolko DJ, Campo JV, Kelleher K, Cheng Y. 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*. 2010;31(5):393.

115. Kolko DJ, Cheng Y, Campo JV, Kelleher K. Moderators and predictors of clinical outcome in a randomized trial for behavior problems in pediatric primary care. *Journal of Pediatric Psychology*. 2011;36(7):753-765.

116. Kolko DJ, Hart JA, Campo J, et al. Effects of Collaborative Care for Comorbid Attention Deficit Hyperactivity Disorder Among Children With Behavior Problems in Pediatric Primary Care. *Clin Pediatr (Phila)*. Jul 2020;59(8):787-800. doi:10.1177/0009922820920013

117. Krägeloh CU, Czuba KJ, Billington DR, Kersten P, Siegert RJ. Using feedback from patient-reported outcome measures in mental health services: a scoping study and typology. *Psychiatric Services*. 2015;66(3):224-241.

118. Kuhlthau K, Jellinek M, White G, VanCleave J, Simons J, Murphy M. Increases in behavioral health screening in pediatric care for Massachusetts Medicaid patients. *Archives of pediatrics & adolescent medicine*. 2011;165(7):660-664.

119. Lambert MJ, Whipple JL, Vermeersch DA, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical psychology & psychotherapy*. 2002;9(2):91-103.

120. Lavigne JV, Meyers KM, Feldman M. Systematic Review: Classification Accuracy of Behavioral Screening Measures for Use in Integrated Primary Care Settings. *Journal of Pediatric Psychology*. 2016;41(10):1091-1109. doi:10.1093/jpepsy/jsw049

121. Lawthers A, Kane M, Sherwood E. Clinical Topic Review 2013-Behavioral Health Screening Among MassHealth Children and Adolescents. 2014;

122. Lee CC, Holder-Niles FF, Haynes L, et al. Associations Between Patient-Reported Outcome Measures of Asthma Control and Psychosocial Symptoms. *Clin Pediatr (Phila)*. 03 2019;58(3):307-312. doi:10.1177/0009922818812479

123. Leiva L, Antivilo-Bruna A, Torres-Cortés B, Peña F, Scquicciarini AM. Relationship between adverse childhood experiences and mental health: Implications for a nationwide school mental health program. *Current Psychology*. 2021/10/13 2021;doi:10.1007/s12144-021-02342-x

124. Leiva L, Rojas R, Peña F, Vargas B, Scquicciarini AM. Detectando las Dificultades Emocionales y Conductuales en la Escuela: Validación de PSC-17. *Revista Iberoamericana de Diagnóstico y Evaluación-e Avaliação Psicológica*. 2019;1(50):95-105.

125. Lin Y-M, Kuo S-Y, Chang Y-K, et al. Effects of Parental Education on Screen Time, Sleep Disturbances, and Psychosocial Adaptation Among Asian Preschoolers: A Randomized Controlled Study. *Journal of Pediatric Nursing-Nursing Care of Children & Families*. Jan-Feb 2021;56:E27-E34. doi:10.1016/j.pedn.2020.07.003

126. Liu J, DiStefano C, Burgess Y, Wang J. Pediatric Symptom Checklist-17 Testing Measurement Invariance of a Higher-Order Factor Model Between Boys and Girls. *European Journal of Psychological Assessment*. Jan 2020;36(1):77-83. doi:10.1027/1015-5759/a000495

127. Liu J, Guo S, Gao R, DiStefano C. Investigating school children’s behavioral and emotional problems using pediatric symptoms checklist-17 in a structural equation modeling framework. *School psychology international*. 2020;41(3):257-275.

128. López V, Murphy M, Lucke C, et al. Peer victimization and mental health risk in Chilean students. *Journal of child and family studies*. 2018;27(8):2608-2621.

129. Lund E, Gabrielli J. The role of pediatric psychologists in mitigating disability-specific barriers among youth during the COVID-19 pandemic. *Clinical Practice in Pediatric Psychology*. 2021;9doi:10.1037/cpp0000387

130. MacNeily AE. Should Psychological Assessment be a Part of Incontinence Management in Children and Adolescents? *The Journal of urology*. 2016;195(5):1327-1328.

131. Mancini MA. A Pilot Study Evaluating a School-Based, Trauma-Focused Intervention for Immigrant and Refugee Youth. *Child and Adolescent Social Work Journal*. Jun 2020;37(3):287-300. doi:10.1007/s10560-019-00641-8

132. Mann C. CMCS Informational Bulletin; Prevention and Early Identification of Mental Health and Substance Use Conditions. 2013;

133. Marlow M, Servili C, Tomlinson M. A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries. *Autism Res*. 02 2019;12(2):176-199. doi:10.1002/aur.2033

134. Marti FA, Pourat N, Lee C, Zima BT. Overview of Ten Child Mental Health Clinical Outcome Measures: Testing of Psychometric Properties with Diverse Client Populations in the US. Article; Early Access. *Adm Policy Ment Health*. 2021;doi:10.1007/s10488-021-01157-z

135. MassHealth. PCC and service location comparison: MGH-Chelsea HealthCare Center Report Card. 2012;Appendix:55.

136. MassHealth. Quarterly Behavioral Health Screening Report: Behavioral health screening January 2008-December 2013. Hingham, MA: Executive Office of Health and Human Services. 2014;

137. May J, Kazee N, Castillo S, et al. From silos to an innovative health care delivery and patient engagement model for children in Medicaid. *Healthcare (Amsterdam)*. 2018;6(1):67-73. doi:10.1016/j.hjdsi.2016.12.008

138. McCarthy A, Asghar S, Wilens T, et al. Using a Brief Parent-Report Measure to Track Outcomes for Children and Teens with ADHD. *Child Psychiatry & Human Development*. 2016;47(3):407-416.

139. McGovern CM, Arcoleo K, Melnyk B. COPE for asthma: Outcomes of a cognitive behavioral intervention for children with asthma and anxiety. *Sch Psychol*. Nov 2019;34(6):665-676. doi:10.1037/spq0000310

140. McLaurin-Jiang S, Cohen GM, Brown CL, Edwards P, Albertini LW. Integrated Mental Health Training Relates to Pediatric Residents' Confidence with Child Mental Health Disorders. *Acad Psychiatry*. Jun 2020;44(3):299-304. doi:10.1007/s40596-020-01182-4

141. Mehus CJ, Kazlauskaite V, Colianni S, Borowsky IW. Parents' interest in resources to address their child's behavioral health through primary care. *Fam Syst Health*. 09 2019;37(3):244-248. doi:10.1037/fsh0000429

142. Mekori-Domachevsky E, Matalon N, Mayer Y, et al. Internalizing symptoms impede adolescents' ability to transition from in-person to online mental health services during the 2019 coronavirus disease pandemic. Article; Early Access. *J Telemed Telecare*. 2021;1357633x211021293. doi:10.1177/1357633x211021293

143. Metwally AM, El-Sonbaty MM, El Etreby LA, et al. Impact of National Egyptian school feeding program on growth, development, and school achievement of school children. *World J Pediatr*. Aug 2020;16(4):393-400. doi:10.1007/s12519-020-00342-8

144. Michaud AL, Cadigan C, Mason C, Recklitis CJ. Evaluating the Brief Parental Intake Form (BPIF) for psychosocial difficulties in childhood cancer survivors. *Psychooncology*. 10 2020;29(10):1510-1512. doi:10.1002/pon.5482

145. Muppidathi S, Boj J, Kunjithapatham M. Use of the pediatric symptom checklist to screen for behaviour problems in children. *International Journal of Contemporary Pediatrics*. 04/25 2017;4:886. doi:10.18203/2349-3291.ijcp20171692

146. Murphy JM, Arnett HL, Bishop SJ, Jellinek MS, Reede JY. Screening for psychosocial dysfunction in pediatric practice: A naturalistic study of the Pediatric Symptom Checklist. *Clinical pediatrics*. 1992;31(11):660-667.

147. Murphy JM, Bergmann P, Chiang C, et al. The PSC-17: subscale scores, reliability, and factor structure in a new national sample. *Pediatrics*. 2016:e20160038.

148. Murphy JM, Blais M, Baer L, et al. Measuring outcomes in outpatient child psychiatry: Reliable improvement, deterioration, and clinically significant improvement. *Clinical child psychology and psychiatry*. 2015;20(1):39-52.

149. Murphy JM, Guzmán J, McCarthy AE, et al. Mental health predicts better academic outcomes: a longitudinal study of elementary school students in Chile. *Child Psychiatry & Human Development*. 2015;46(2):245-256.

150. Murphy JM, Ichinose C, Hicks RC, et al. Utility of the Pediatric Symptom Checklist as a psychosocial screen to meet the federal Early and Periodic Screening, Diagnosis, and Treatment (EPSDT) standards: a pilot study. *The Journal of Pediatrics*. 1996;129(6):864-869.

151. Murphy JM, Jellinek M, Milinsky S. The Pediatric Symptom Checklist: validation in the real world of middle school. *Journal of Pediatric Psychology*. 1989;14(4):629-639.

152. Murphy JM, Masek B, Babcock R, et al. Measuring outcomes in outpatient child psychiatry: The contribution of electronic technologies and parent report. *Clinical child psychology and psychiatry*. 2011;16(1):146-160.

153. Murphy JM, Nguyen T, Lucke C, Chiang C, Plasencia N, Jellinek M. Adolescent self-screening for mental health problems; Demonstration of an internet-based approach. *Academic Pediatrics*. 2017;

154. Murphy JM, Nguyen T, Lucke C, Chiang C, Plasencia N, Jellinek M. Adolescent Self-Screening for Mental Health Problems; Demonstration of an Internet-Based Approach. *Acad Pediatr*. 2018 Jan - Feb 2018;18(1):59-65. doi:10.1016/j.acap.2017.08.013

155. Murphy JM, Reede J, Jellinek MS, Bishop SJ. Screening for psychosocial dysfunction in inner-city children: further validation of the Pediatric Symptom Checklist. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1992;31(6):1105-1111.

156. Murphy JM, Riobueno-Naylor A, Haile H, et al. Behavioural Health Screening and Service Use in a Statewide Sample of Medicaid-Eligible Pediatric Outpatients. *Science Repository*. 2020;doi:10.31487/j.PDR.2020.03.04

157. Murphy JM, Stepanian S, Riobueno-Naylor A, et al. Implementation of an electronic approach to psychosocial screening in a network of pediatric practices. *Academic pediatrics*. 2021;21(4):702-709.

158. Murphy KP, Rashleigh CM, Timulak L. The relationship between progress feedback and therapeutic outcome in student counselling: A randomised control trial. *Counselling Psychology Quarterly*. 2012;25(1):1-18.

159. Murphy M, Kamin H, Masek B, et al. Using brief clinician and parent measures to track outcomes in outpatient child psychiatry: longer term follow-up and comparative effectiveness. *Child and Adolescent Mental Health*. 2012;17(4):222-230.

160. Murray CJ, Lopez AD. *Global burden of disease*. vol 1. Harvard University Press Boston; 1996.

161. Nasir A, Zimmer A, Taylor D, Santo J. Psychosocial assessment of the family in the clinical setting. *BMC Psychol*. Jan 11 2019;7(1):3. doi:10.1186/s40359-018-0277-5

162. O'Connell, Boat, Warner. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. Washington, D.C.: Committee on the Prevention of Mental Disorders and Substance Abuse Among Children Youth and Young Adults. 2009;

163. Oliva S, Russo G, Gili R, et al. Risks and Protective Factors Associated With Mental Health Symptoms During COVID-19 Home Confinement in Italian Children and Adolescents: The# Understandingkids Study. *Frontiers in Pediatrics*. 2021;9:537.

164. Omkarappa DB, Rentala S, Nattala P. Psychiatric nurse delivered group-cognitive-behavioral therapy for internalizing behavior problems among children of parents with alcohol use disorders. *Journal of Child and Adolescent Psychiatric Nursing*. 2021;

165. Öner Ö, Karaca S, Cangür Ş. The validity and reliability of pictorial pediatric symptom checklist in children of 4-5 ages. *Anadolu Psikiyatri Derg*. 2019;20(1):73-81.

166. Ozturk FO, Ayaz-Alkaya S. Internet addiction and psychosocial problems among adolescents during the COVID-19 pandemic: A cross-sectional study. Article. *Arch Psychiatr Nurs*. Oct 2021;35(5):595-601. doi:10.1016/j.apnu.2021.08.007

167. Pagano ME, Cassidy LJ, Little M, Murphy JM, Jellinek MS. Identifying psychosocial dysfunction in school-age children: The Pediatric Symptom Checklist as a self-report measure. *Psychology in the Schools*. 2000;37(2):91.

168. Parker EM, Jacobson J, Pullmann MD, Kerns SEU. Identifying Psychosocial Problems Among Children and Youth in the Child Welfare System Using the PSC-17: Exploring Convergent and Discriminant Validity with Multiple Informants. *Child Psychiatry Hum Dev*. 02 2019;50(1):108-120. doi:10.1007/s10578-018-0824-6

169. Pesek R, Whitaker B, Ararat E, et al. Pediatric Severe Asthmatics in a Southern US State Face Multiple Challenges and Barriers to Care. Meeting Abstract. *J Allergy Clin Immunol*. Feb 2021;147(2):AB46-AB46.

170. Pourat N, Zima B, Marti A, Lee C. California child mental health performance outcomes system: Recommendation report. *UCLA Center for Health Policy Research*. 2017;

171. Raman V, Prakash A, D'Souza F. Psychosocial Issues in Children With Thalassemia: From Identification to a Model for Management in a Developing Country. *J Pediatr Hematol Oncol*. 04 2019;41(3):218-221. doi:10.1097/mph.0000000000001370

172. Ratcliff MB, Catlin PA, Peugh JL, Siegel RM, Kirk S, Tamm L. Psychosocial Screening Among Youth Seeking Weight Management Treatment. *Clin Pediatr (Phila)*. 03 2018;57(3):277-284. doi:10.1177/0009922817715936

173. Reese RJ, Norsworthy LA, Rowlands SR. Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, research, practice, training*. 2009;46(4):418.

174. Reese RJ, Toland MD, Slone NC, Norsworthy LA. Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy: Theory, research, practice, training*. 2010;47(4):616.

175. Reijneveld SA, Vogels AG, Hoekstra F, Crone MR. Use of the Pediatric Symptom Checklist for the detection of psychosocial problems in preventive child healthcare. *BMC Public Health*. 2006;6(1):197.

176. Richard HM, Cerza SP, De La Rocha A, Podeszwa DA. Preoperative mental health status is a significant predictor of postoperative outcomes in adolescents treated with hip preservation surgery. *Journal of Childrens Orthopaedics*. 2020;14(4):259-265. doi:https://doi.org/10.1302/1863-2548.14.200013

177. Richardson LP, Ludman E, McCauley E, et al. Collaborative care for adolescents with depression in primary care: a randomized clinical trial. *Jama*. 2014;312(8):809-816.

178. Richardson LP, McCauley E, Grossman DC, et al. Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. *Pediatrics*. 2010;126(6):1117-1123.

179. Riffat O, Humayun Iqbal K, Muhammad Khalid M, Najaf M, Fatima T. Psychosocial impact of the COVID-19 pandemic on doctors' children: are we heading towards a mental health pandemic? article. *Paediatrica Indonesiana*. 02/01/ 2021;61(1):46-52. doi:10.14238/pi61.1.2021.46-52

180. Riobueno-Naylor A, Romo S, Kazis L, et al. Usefulness of a Novel System for Feedback of Patient-Reported Outcome Measures in Children Recovering From Burns. *J Burn Care Res*. 10 16 2019;40(6):776-784. doi:10.1093/jbcr/irz082

181. Riobueno-Naylor A, Williamson H, Canenguez K, et al. Appearance Concerns, Psychosocial Outcomes, and the Feasibility of Implementing an Online Intervention for Adolescents Receiving Outpatient Burn Care. *J Burn Care Res*. 02 03 2021;42(1):32-40. doi:10.1093/jbcr/iraa108

182. Roby E, Piccolo LR, Gutierrez J, et al. Father involvement in infancy predicts behavior and response to chronic stress in middle childhood in a low-income Latinx sample. *Developmental Psychobiology*. 2021;doi:10.1002/dev.22081

183. Romano-Clarke G, Tang MH, Xerras DC, et al. Have rates of behavioral health assessment and treatment increased for Massachusetts children since the Rosie D. decision? A report from two primary care practices. *Clinical pediatrics*. 2014;53(3):243-249.

184. Rukabyarwema JP, McCall N, Ngambe T, Kanyembari XB, Needlman R. Behavior Problems in Physically Ill Children in Rwanda. *Journal of developmental and behavioral pediatrics : JDBP*. 2019 Oct/Nov 2019;40(8):642-650. doi:10.1097/dbp.0000000000000698

185. Rushton J, Bruckman D, Kelleher K. Primary Care Referral of Children With Psychosocial Problems. *Year Book of Psychiatry & Applied Mental Health*. 2004;2004(1):80-81.

186. Sachs GS, Thase ME, Otto MW, et al. Rationale, design, and methods of the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological psychiatry*. 2003;53(11):1028-1042.

187. Sachser C, Berliner L, Holt T, et al. International development and psychometric properties of the Child and Adolescent Trauma Screen (CATS). *J Affect Disord*. Mar 01 2017;210:189-195. doi:10.1016/j.jad.2016.12.040

188. Savageau JA, Simons, J., Lucke, C., Jellinek, M., Murphy, J.M. Assessing Disparities in Mental Health Screening and Services Before and After Implementation of an Innovative Statewide Program for Children with Medicaid. 2017:

189. Savageau JA, Keller D, Willis G, et al. Behavioral Health Screening among Massachusetts Children Receiving Medicaid. *The Journal of Pediatrics*. 2016;178:261-267.

190. Semansky R, Koyanagi C, Vandivort-Warren R. Behavioral health screening policies in Medicaid programs nationwide. *Psychiatric Services*. 2003;54(5):736-739.

191. Semansky RM, Koyanagi C, Vandivort-Warren R. Behavioral health screening policies in Medicaid programs nationwide. *Psychiatric Services*. 2003;54(5):736-739.

192. Semenkovich K, Berlin KS, Ankney RL, et al. Predictors of diabetic ketoacidosis hospitalizations and hemoglobin A1c among youth with Type 1 diabetes. *Health Psychol*. Jul 2019;38(7):577-585. doi:10.1037/hea0000719

193. Shaffer A, Lindhiem O, Kolko D. Treatment Effects of a Primary Care Intervention on Parenting Behaviors: Sometimes It's Relative. *Prev Sci*. 04 2017;18(3):305-311. doi:10.1007/s11121-016-0689-5

194. Shah S, Kaul A, Shah R, Maddipoti S. Impact of Coronavirus Disease 2019 Pandemic and Lockdown on Mental Health Symptoms in Children. *Indian Pediatrics*. Jan 2021;58(1):75-76.

195. Shahzad S, Kliewer W, Ali M, Begum N. Different risk factors are associated with alcohol use versus problematic use in male Pakistani adolescents. *International Journal of Psychology*. Aug 2020;55(4):585-589. doi:10.1002/ijop.12617

196. Shellman AB, Meinert AC, Curtis DF. Physician Utilization of a Universal Psychosocial Screening Protocol in Pediatric Primary Care. *Clin Pediatr (Phila)*. 08 2019;58(9):957-969. doi:10.1177/0009922819845878

197. Simha S, Brown AC. Preventive Care in Children and Adolescents. *Primary Care: Clinics in Office Practice*. 2021;48(1):99-116.

198. Simon W, Lambert MJ, Harris MW, Busath G, Vazquez A. Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy research*. 2012;22(6):638-647.

199. Siu AL. Screening for depression in children and adolescents: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*. 2016;164(5):360-366.

200. Slade K, Lambert MJ, Harmon SC, Smart DW, Bailey R. Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical psychology & psychotherapy*. 2008;15(5):287-303.

201. Spencer AE, Baul TD, Sikov J, et al. The relationship between social risks and the mental health of school-age children in primary care. *Academic pediatrics*. 2020;20(2):208-215.

202. Spencer AE, Chiang C, Plasencia N, et al. Utilization of Child Psychiatry Consultation Embedded in Primary Care for an Urban, Latino Population. *Journal of Health Care for the Poor & Underserved*. 2019;30(2):637-652. doi:10.1353/hpu.2019.0047

203. Spencer AE, Plasencia N, Sun Y, et al. Screening for Attention-Deficit/Hyperactivity Disorder and Comorbidities in a Diverse, Urban Primary Care Setting. *Clin Pediatr (Phila)*. 10 2018;57(12):1442-1452. doi:10.1177/0009922818787329

204. Stadnick NA, Martinez K, Aarons GA, Lee DA, Van Cleave J, Brookman-Frazee L. Pediatric Primary Care Perspectives on Integrated Mental Health Care for Autism. *Academic pediatrics*. 2020/11/01/ 2020;20(8):1140-1147. doi:https://doi.org/10.1016/j.acap.2020.03.006

205. Stein BD, Jaycox LH, Kataoka SH, et al. A mental health intervention for schoolchildren exposed to violence: A randomized controlled trial. *Jama*. 2003;290(5):603-611.

206. Struemph KL, Barhight LR, Thacker D, Sood E. Systematic psychosocial screening in a paediatric cardiology clinic: clinical utility of the Pediatric Symptom Checklist 17. *Cardiology in the Young*. 2016;26(06):1130-1136.

207. Stubbs PH, Walters AS. Tools for the Assessment of Pediatric Restless Legs Syndrome. *Front Psychiatry*. 2020;11:356. doi:10.3389/fpsyt.2020.00356

208. Studts CR, Polaha J, van Zyl MA. Identifying Unbiased Items for Screening Preschoolers for Disruptive Behavior Problems. *J Pediatr Psychol*. 05 01 2017;42(4):476-486. doi:10.1093/jpepsy/jsw090

209. Suku S, Soni J, Martin MA, et al. A multivariable analysis of childhood psychosocial behaviour and household functionality. *Child Care Health Dev*. 07 2019;45(4):551-558. doi:10.1111/cch.12665

210. Terrone G, Parente I, Romano A, Auricchio R, Greco L, Del Giudice E. The Pediatric Symptom Checklist as screening tool for neurological and psychosocial problems in a paediatric cohort of patients with coeliac disease. *Acta Paediatrica*. 2013;102(7):e325-e328.

211. Thomaseo Burton E, Jones TL, Smith WA, Han JC. Psychosocial Functioning After One Year of Interdisciplinary Pediatric Weight Management. *Behav Med*. 2020 Apr-Jun 2020;46(2):92-99. doi:10.1080/08964289.2019.1570073

212. Tolliver M, Polaha J, Williams SL, Studts CR. Evaluating the Prevalence of Child Psychosocial Concerns in Rural Primary Care. *Rural Ment Health*. Apr 2020;44(2):96-105. doi:10.1037/rmh0000130

213. Trafalis S, Giannini C, Joves J, et al. A pediatrician-friendly review of three common behavioral health screeners in pediatric practice: Findings and recommendations. *Pediatric Investigation*. Mar 2021;5(1):58-64. doi:10.1002/ped4.12246

214. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR* D: implications for clinical practice. *American journal of Psychiatry*. 2006;163(1):28-40.

215. Vargas B, Leiva L, Rojas-Andrade R, Scquicciarini AM. Effects of Psychosocial Adversity on School Maladjustment: A Follow-up of Primary School Students. article. *Psicología Educativa: Revista de los Psicólogos de la Educación*. 2019;25(2):101-108. doi:10.5093/psed2019a10

216. Weigel MM, Armijos RX. Household Food Insecurity and Psychosocial Dysfunction in Ecuadorian Elementary Schoolchildren. *Int J Pediatr*. 2018;2018:6067283. doi:10.1155/2018/6067283

217. Weitzman C, Wegner L. Promoting optimal development: Screening for behavioral and emotional problems. *Pediatrics*. 2015:peds. 2014-3716.

218. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology*. 2003;50(1):59.

219. Wissow L, Brown J, Fothergill K, et al. Universal Mental Health Screening in Pediatric Primary Care: A Systematic Review. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2013;52(11):1134 - 1147.e23.

220. Wolfe-Christensen C, Guy WC, Mancini M, Kovacevic LG, Lakshmanan Y. Evidence of need to use self-report measures of psychosocial functioning in older children and adolescents with voiding dysfunction. *The Journal of urology*. 2016;195(5):1570-1574.

221. Wylie-Rosett J, Groisman-Perelstein AE, Diamantis PM, et al. Embedding weight management into safety-net pediatric primary care: randomized controlled trial. *Int J Behav Nutr Phys Act*. 01 22 2018;15(1):12. doi:10.1186/s12966-017-0639-z

222. Young ND, Takala CR. Sequential Screening to Improve Behavioral Health Needs Detection in Primary Care. *J Am Acad Child Adolesc Psychiatry*. 08 2018;57(8):603-609. doi:10.1016/j.jaac.2018.04.017

223. Zalta AK, Bui E, Karnik NS, et al. Examining the Relationship Between Parent and Child Psychopathology in Treatment-Seeking Veterans. *Child Psychiatry Hum Dev*. 04 2018;49(2):209-216. doi:10.1007/s10578-017-0743-y

224. Zima BT, Marti FA, Lee CE, Pourat N. Selection of a Child Clinical Outcome Measure for Statewide Use in Publicly Funded Outpatient Mental Health Programs. *Psychiatr Serv*. 05 01 2019;70(5):381-388. doi:10.1176/appi.ps.201800424

225. Zimmerman M, McGlinchey JB. Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. *Annals of Clinical Psychiatry*. 2008;20(3):125-129.

Contact Information

Measure Steward (Intellectual Property Owner) : Massachusetts General Hospital
Measure Steward Point of Contact: Wilens, Timothy, twilens@mgh.harvard.edu
Measure Developer if different from Measure Steward: Massachusetts General Hospital
Measure Developer Point(s) of Contact: Murphy, Michael, mmurphy6@partners.org

Additional Information

1. Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.

[Response Begins]
Available in attached file
[Response Ends]

Attachment: 3332_Full NQF Appendix_11-2-21_posted 11-2-21.docx

2. List the workgroup/panel members' names and organizations.

Describe the members' role in measure development.

[Response Begins]
N/A
[Response Ends]

3. Indicate the year the measure was first released.

[Response Begins]
1984
[Response Ends]

4. Indicate the month and year of the most recent revision.

[Response Begins]
January 2020
[Response Ends]

5. Indicate the frequency of review, or an update schedule, for this measure.

[Response Begins]
continuous/ongoing
[Response Ends]

6. Indicate the next scheduled update or review of this measure.

[Response Begins]
Continuous/ongoing
[Response Ends]

7. Provide a copyright statement, if applicable. Otherwise, indicate "N/A".

[Response Begins]
©1988, M.S. Jellinek and J.M. Murphy, Massachusetts General Hospital
[Response Ends]

8. State any disclaimers, if applicable. Otherwise, indicate "N/A".

[Response Begins]
None
[Response Ends]

9. Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".

[Response Begins]
N/A
[Response Ends]