Understanding Developmental-Behavioral Screening Measures
Kevin P. Marks and Angela C. LaRosa
Pediatrics in Review 2012;33;448
DOI: 10.1542/pir.33-10-448

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pedsinreview.aappublications.org/content/33/10/448

Data Supplement (unedited) at:
http://pedsinreview.aappublications.org/content/suppl/2012/09/17/33.10.448.DCSupplementary_Data.html
http://pedsinreview.aappublications.org/content/suppl/2012/09/26/33.10.448.DC1.html
http://pedsinreview.aappublications.org/content/suppl/2012/09/26/33.10.448.DC2.html
Understanding Developmental–Behavioral Screening Measures

Kevin P. Marks, MD,* †
Angela C. LaRosa, MD ‡

Educational Gap

With nearly half of pediatricians still not using standardized screening tools, primary care clinicians should know that the periodic use of general developmental and social–emotional screening tools has been proven to identify promptly two to six times more children (age 0–5 years) with suspected delays than a clinician’s unstructured surveillance alone.

Objectives After completing this article, readers should be able to:

1. Recognize the many methods available to detect a wide spectrum of developmental and behavioral conditions with screening tools, and the screening schedule recommended by the AAP.
2. Recognize that universal, periodic screening with psychometrically sound developmental–behavioral instruments greatly enhances a clinician’s ability to detect, refer, and monitor children who have evolving developmental and behavioral problems. A pediatric provider’s impression has good specificity but poor sensitivity and thus has difficulty identifying the less obvious delays.
3. Know that children who have false-positive screens (screening test concerning but subsequently ineligible for early intervention or early childhood special education services) are a high-risk group in need of close monitoring and other community services such as evidence-based parenting programs and Head Start.
4. Know the main limitations of applying standardized screening tools into a primary care setting.
5. Be aware that screening measures which adhere to core psychometric and feasibility standards are more likely to perform consistently well and be found sustainable in a primary care setting.

Introduction

Developmental and behavioral problems are among the most common conditions of childhood and adolescence; 15% of children have a developmental disability and 21% have a mental health disorder. If intervention is instituted before kindergarten entrance, many problems can be prevented and the large majority can be ameliorated. In the United States, early detection depends heavily on primary care providers. Primary care clinicians should know that the periodic use of general developmental (1)(2)(3)(4) and social–emotional (5) screening tools has been proven to identify promptly two to six times more children (age 0–5 years) who have suspected delays than a clinician’s surveillance alone. (1)(2)(3)(4)(5) This difference in outcome includes foster care populations, in whom pediatricians possess knowledge of children being at higher risk for a developmental delay or social–emotional disorder. (4)(5)

Abbreviations

ASQ: Ages and Stages Questionnaire
ASQ-3: Ages & Stages Questionnaire, Third Edition
ASD: autism spectrum disorder
ECSE: early childhood special education
EI: early intervention
M–CHAT: Modified Checklist for Autism in Toddlers
NPV: negative predictive value
PEDS: Parents’ Evaluation of Developmental Status
PPV: positive predictive value

*Department of Pediatrics, PeaceHealth Medical Group, Eugene, OR.
†Division of General Pediatrics, School of Medicine, Oregon Health & Science University, Portland, OR.
‡Department of Genetics and Developmental–Behavioral Pediatrics, Medical University South Carolina, Charleston, SC.

Downloaded from http://pedsinreview.aappublications.org/ by Kevin Marks on October 1, 2012
Periodic screening leads to substantial increases in early intervention (EI) and early childhood special education (ECSE) eligibility rates. (1)(2) Interestingly, the percentage of pediatricians “who self-reported always/almost always using ≥1 screening tools” has increased significantly from 23.0% to 47.7% from 2002 to 2009. (6) With greater numbers of clinicians incorporating standardized screening tools into their practices, a better understanding of how to optimize their performance is needed. With nearly half of pediatricians still not routinely using standardized tools, further education is warranted about which tools are feasible for use in primary care, with attention paid to which tools are psychometrically sound. Education also is warranted on how practitioners can better employ screens in a 1) safe, 2) equitable, 3) effective, 4) timely, 5) patient- and parent-centered, and 6) efficient manner. These are the Institute of Medicine’s six quality aims.

Surveillance is the process in which clinicians watch for signs of developmental or behavioral problems in the course of caring for children. Screening refers to the use of brief standardized tools to differentiate those who need a further evaluation for potential problems from those who probably do not. Although screening measures help to refine the pattern of delays in children who have detectable problems, they are meant primarily for children perceived to be asymptomatic by the clinician and for those in whom surveillance demonstrates risk.

Developmental-behavioral screens are different from many other medical screens, which seek to identify a strictly positive or negative condition (eg, cystic fibrosis, congenital hypothyroidism, phenylketonuria). In fact, research indicates that children who have false-positive screens (screening test positive/concerning but subsequently ineligible for EI/ECSE services) are a high-risk group in need of diligent monitoring and other community resources. (2)(7)

There are numerous categories of developmental-behavioral screening instruments (Table 1) used to identify a wide spectrum of developmental-behavioral conditions with variable prevalence rates at different ages (Table 2). Parent-report screens (Figs 1–3) typically are designed to measure one of the first three components of developmental-behavioral surveillance: 1) eliciting and addressing parents’ concerns; 2) maintaining a developmental-behavioral history (reviewing milestones and behaviors); and 3) identifying developmental-behavioral risk and protective factors (biological and environmental).

Table 1. Common Developmental–Behavioral Screening Tool Categories

<table>
<thead>
<tr>
<th>I. Parent-report: measures parental concerns versus; milestones/skills and behaviors versus; developmental–behavioral (biological and environmental) risk and/or protective factors versus; a history of parent–child interactions versus a combination of all of the above.</th>
<th>II. Practitioner-administered: directly eliciting and/or observing milestones/skills and behaviors about the child and/or the parent–child interactions. In reality, many practitioner-administered tools have parent-report items.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Broad-band Screening Measures</strong></td>
<td><strong>Narrow-band Screening Measures</strong></td>
</tr>
</tbody>
</table>
| A. Broad-band family psychosocial | 1. Parental mood disorder/depression-specific
2. Parental substance abuse-specific
3. Domestic violence–specific
4. Parental adverse childhood experience–specific |
| B. Broad-band (general) developmental | 1. Social–emotional
2. Language- and cognitive-specific
3. Communication–specific
4. Autism–specific
5. Motor–specific |
| C. Broad-band mental health/behavioral/psychosocial function | 1. ADHD–, oppositional defiant–, conduct disorder–specific
2. Academic– and/or literacy–specific
3. Anxiety disorder–specific
4. Depression/mood disorder–specific
5. Substance abuse–specific
6. Suicide–specific |

ADHD = attention-deficit/hyperactivity disorder.

*aThe AAP recommends universal postpartum maternal mood disorder screening in the first year (experts recommend 2, 4, and/or 6 months), general developmental screening at 9, 18, and 30 (or 24) months (AAP experts also recommend a “kindergarten readiness” screening at 4 years), autism-specific screening at 18 and 24 months, mental health/psychosocial function screening at every health supervision visit ≥5 years, and substance abuse–specific screening at every health supervision visit throughout adolescence. (6)

*bThe US Prevention Service Task Force recommends depression-specific screening at every health supervision visit throughout adolescence.
Practitioner-administered screens (Fig 4), which directly elicit or make observations about the child’s skills and behavior, generally measure another component of surveillance, making accurate and informed observations about the child and parent–child interactions. Screens may be broad-band (general) (Figs 1 and 4), meaning they tap all or most developmental-behavioral domains (expressive speech and language, receptive language, gross motor, fine motor, cognitive/problem-solving, self-help and adaptive skills, social-emotional/behavior) or narrow-band (Fig 2), meaning domain-, disorder-, or disability-specific. Narrow-band screens may partially tap one or more domains but are not designed to assess the wide spectrum of development-behavioral conditions at various ages (Table 2) and, therefore, ideally should always be used in conjunction with a broad-band screen.

Current Recommendations

The American Academy of Pediatrics (AAP) currently recommends universal postpartum mood disorder screening in the first year after birth; general developmental screening at 9, 18, and 30 (or 24) months; autism-specific screening at 18 and 24 months; social-emotional screening whenever a general or autism-specific instrument is abnormal; “kindergarten readiness” screening at 4 years; social-emotional/mental health/psychosocial function screening at every health supervision visit from ages 5 to 18 years; and substance abuse–specific screening at every health supervision visit throughout adolescence.

In addition, according to the AAP, an appropriate screen should be administered whenever a clinician’s surveillance determines “risk.” All of these screening recommendations have heightened the need for clinicians to implement office-based and community-based systems that are simultaneously effective and efficient, especially for those for whom research suggests interventions are most effective: the 25.7 million children in the United States who are age <5 years (projected population in 2012 per http://www.childstats.gov).

Incorporating Periodic Screening Into Ongoing Surveillance

Figure 5 provides a collaborative model for screening and surveillance in a medical home setting and is a combination of the 2006 developmental, 2007 autism spectrum disorders (ASDs), and 2010 mental health AAP

---

Table 2. Prevalence of Developmental–Behavioral Conditions in Childrena

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prevalence (%; most studies conducted in the United States)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any developmental disability</td>
<td>15.0% (3–17 y in 2006–2008)</td>
</tr>
<tr>
<td>1. Social–emotional/behavioral disorders</td>
<td>21% (children and adolescents who meet criteria for a mental health disorder with at least minimal impairment in 1999)</td>
</tr>
<tr>
<td>2. Speech and language impairment</td>
<td>9.5%–14.2% (“serious emotional disturbance,” 0–5 y in 2006)</td>
</tr>
<tr>
<td>3. Attention–deficit/hyperactivity disorder</td>
<td>13.5% (expressive speech delays, 18–23 mo in 1996–2000)</td>
</tr>
<tr>
<td>4. Learning disability</td>
<td>17.5% (expressive speech delays, 30–36 mo in 1996–2000)</td>
</tr>
<tr>
<td>5. Developmental coordination disorder</td>
<td>7.4% (language impairment, kindergarten in 1997)</td>
</tr>
<tr>
<td>7. ASDs</td>
<td>4.1% (4–8 y in 2006–2008)</td>
</tr>
<tr>
<td>8. Cerebral palsy</td>
<td>7.2% (3–17 y in 2006–2008)</td>
</tr>
<tr>
<td>9. Hearing impairmentc</td>
<td>0.67% (3–17 y in 2006–2008)</td>
</tr>
<tr>
<td>10. Vision impairmentd</td>
<td>0.66% (8 y and 3–10 y in 1998)</td>
</tr>
<tr>
<td>Less common conditions (&lt; 1%)</td>
<td>1.1%b (3–17 y in 2006–2008)</td>
</tr>
</tbody>
</table>

ASD = autism spectrum disorder.

aPlease note that all of these developmental–behavioral conditions have high rates of comorbidity.

bIn 2007, the ASD point prevalence was 110 per 10,000 (1.1%). Diagnosis of ASD was defined by 1) a physician or other health-care provider who said that the child had ASD and 2) that the child currently had the condition. Nearly 40% of those ever diagnosed with ASD did not currently have the condition at the time of a random-digit-dial telephone National Survey of Children’s Health (sample size = 78,037).

cDefined by the Centers for Disease Control and Prevention as moderate hearing loss (40 dB) or worse, unaided, in the better ear.

dDefined by the Centers for Disease Control and Prevention as 20/70 visual acuity or worse, with correction, in the better eye.
algorithms. (8) When a psychometrically sound screen raises concerns, the algorithm emphasizes system-wide referral care coordination and a secondary screening or assessment, which, in a primary care setting, is best accomplished with the help of an EI/ECSE agency and an early return (≤1 month) of a face visit.

In this algorithm, pediatric clinicians are encouraged to refer automatically when their clinical impression confidently detects a developmental delay or if the child possesses a biological or environmental condition associated with a sufficiently high probability of a delay; developmental-behavioral promotion has been formally incorporated as a component of surveillance to make the process safer, more effective, and parent-centered; plus greater detail has been provided about the decisions and action steps that should reliably occur after a screening test is administered (ie, postscreening surveillance). (8) “Structured surveillance” means that the process and decision-making are enhanced with periodic screening by using evidence-based measures. (8) “Unstructured surveillance” means that the process and decision-making are reliant on subjective impressions or casual observations. (8) In a systematic review, a pediatric clinician’s unstructured surveillance, when compared with a validated screening tool or a diagnostic interview, has good specificity (the proportion of children correctly identified as not having a developmental-behavioral problem) ranging from 69% to 100%, but poor sensitivity (the proportion of children correctly identified as having a developmental-behavioral problem) ranging from 14% to 54%. (9)

When pediatric health-care providers detect a problem, they usually are correct; however, they struggle to identify the majority of delays, most likely because they over-rely on psychometrically unsound milestone checklists with vague referral criteria. (1)(2)(8) As a result, the large majority of children who have evolving developmental-behavioral problems can be deprived of the benefits of EI, ECSE, or other evidence-based community services before kindergarten entrance.

Periodic screening (ie, structured surveillance) has been proven repeatedly to enhance a clinician’s ability to detect, refer, and monitor children who have evolving developmental-behavioral problems. (1)(2)(3)(4)(5)(8) Limitations of Screening in Primary Care

Although standardized screening improves the quality of care at health supervision visits, increases parental satisfaction, leads to substantial cost savings for society, and positively affects the lives of children and families via the positive developmental outcomes and preventive health care, it has limitations. (1)(2)(3)(4)(5)(8)

<table>
<thead>
<tr>
<th>Psychometric Properties</th>
<th>ASQ-3</th>
<th>IDI &amp; CDR-POQ</th>
<th>POG II</th>
<th>PEDI</th>
<th>PEDI-DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of measure</td>
<td>Parent-report or self-reported (or via interview)</td>
<td>Parent-report or self-reported (or via interview)</td>
<td>Parent-report or self-reported (or via interview)</td>
<td>Parent-report or self-reported (or via interview)</td>
<td>Parent-report or self-reported (or via interview)</td>
</tr>
<tr>
<td>Scores produced</td>
<td>Age-specific cutoffs (topical referents) in 5 developmental domains</td>
<td>Age-specific cutoffs (topical referents) in 5 developmental domains</td>
<td>Age-specific cutoffs (topical referents) in 5 developmental domains</td>
<td>Age-specific cutoffs (topical referents) in 5 developmental domains</td>
<td>Age-specific cutoffs (topical referents) in 5 developmental domains</td>
</tr>
<tr>
<td>National (USA): a diverse sample</td>
<td>yes &amp; yes diverse sample but not nationwide representative of U.S. popul.</td>
<td>questionable</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>General: naturalistic sample</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Standardization sample size &amp; children per age interval is preferred</td>
<td>(200 children per age interval is preferred)</td>
<td>(200 children per age interval is preferred)</td>
<td>(200 children per age interval is preferred)</td>
<td>(200 children per age interval is preferred)</td>
<td>(200 children per age interval is preferred)</td>
</tr>
<tr>
<td>Interval reliability (% CICV)</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>Test-retest reliability (% CICV)</td>
<td>75 - 95%</td>
<td>not available</td>
<td>93%</td>
<td>94%</td>
<td>95%</td>
</tr>
<tr>
<td>Internal consistency (CICV)</td>
<td>90 - 92%</td>
<td>not available</td>
<td>90 - 92%</td>
<td>90 - 92%</td>
<td>90 - 92%</td>
</tr>
<tr>
<td>Appropriate: evidence standards</td>
<td>acceptable</td>
<td>acceptable</td>
<td>not acceptable</td>
<td>acceptable</td>
<td>acceptable</td>
</tr>
<tr>
<td>Noteworthy forms of validity bias</td>
<td>small sample</td>
<td>small sample</td>
<td>small sample</td>
<td>small sample</td>
<td>small sample</td>
</tr>
<tr>
<td>Concurrent criterion-related validity sensitivity score (standardized z)</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Concurrent criterion-related validity specificity score (standardized z)</td>
<td>77%</td>
<td>77%</td>
<td>77%</td>
<td>77%</td>
<td>77%</td>
</tr>
<tr>
<td>Construct, convergent and/or discriminate validity studies</td>
<td>yes &amp; supportive</td>
<td>no</td>
<td>no</td>
<td>yes &amp; supportive</td>
<td>no</td>
</tr>
</tbody>
</table>

Figure 1. Broad-band (general) developmental: parent-report screening measures. Core standards are highlighted in yellow. See text to explain shading and see Table 3 for abbreviations.
benefits of EI and ECSE, clinicians should be aware that screening has its limitations. No single instrument will suffice to identify every developmental-behavioral problem and match the needs of every population or practice setting. The population screened and the method of implementation affect the psychometric properties of a screening measure. Screens can be falsely negative (ie, they can yield negative or borderline results, but the child could be promptly deemed eligible for EI/ECSE services) or falsely positive (ie, the screening test is positive or concerning, but the child is subsequently deemed EI/ECSE ineligible). However, a clinician’s longitudinal surveillance can be used to better interpret the typical or borderline results of a screening test.

For example, when a pediatrician confidently suspects a delay and, simultaneously, a screen such as the Ages and Stages Questionnaire (ASQ) is found to be typical or borderline, 92% of pediatrician-referred children (blind to ASQ results) promptly receive some form of EI services (40% are found to be EI-eligible; 52% are placed on an EI-monitoring list). (1) However, there is no evidence to support the concept that unstructured surveillance should be used to override the concerning results of a psychometrically sound screen and negate the need for a community-based referral. (8) Children who fail screening tests and subsequently are not promptly found to have a developmental-behavioral diagnosis (false-positives) nonetheless tend to have numerous psychosocial risk factors and exhibit performance well below average in the better predictors of school success: pre-academics, language, and intelligence. (7) Indeed, children found to have false-positive screening results are a high-risk group in need of closer monitoring, repeat screenings at subsequent visits, and other community resources (eg, evidence-based parenting programs, nurse home visit programs, high-quality child care, structured preschools such as Head Start).

Evaluating Specific Characteristics of Screening Measures
Although experts agree that no single screening instrument will suffice to identify every developmental-behavioral problem in every practice setting, it is also true that all screening tools are not constructed equal. In 2008, Drotar et al (10) provided guidelines to help pediatricians select which screening tools were best-suited to their practices based on their population’s risk level. Figure 2. Narrow-band parent-report: social-emotional (SE), communication & autism-specific screens. Data not included for the M-CHAT Follow-up Interview (FI), which is recommended for any failed M-CHAT. Core standards are highlighted in yellow. See text to explain shading and see Table 3 for abbreviations.

![Table](https://example.com/table.png)

Figure 2. Narrow-band parent-report: social-emotional (SE), communication & autism-specific screens. Data not included for the M-CHAT Follow-up Interview (FI), which is recommended for any failed M-CHAT. Core standards are highlighted in yellow. See text to explain shading and see Table 3 for abbreviations.
For general developmental parent-report tools, Drotar et al recommended the ASQ and Parents’ Evaluation of Developmental Status (PEDS) for general primary care populations and the ASQ specifically for high-risk populations. For ASD-specific screening tools, the Modified Checklist for Autism in Toddlers (M-CHAT) was recommended for general primary care populations. For further details and recommendations, refer to http://www.cmwf.org/Content/Publications/Fund-Manuals/2008/Feb/Pediatric-Developmental-Screening–Understanding-and-Selecting-Screening-Instruments.aspx.

Since 2008, new tools have been developed (eg, Parents’ Evaluation of Developmental Status–Developmental Milestones [PEDS:DM], Quantitative Checklist for Autism in Toddlers), and others have undergone revisions (eg, Ages & Stages Questionnaire, Third Edition [ASQ-3]). In 2009, an AAP periodic survey, in concert with the screening tool selection trends in national and statewide implementation initiatives, collectively suggested that the ASQ, PEDS, and M-CHAT have emerged as the three most commonly selected screening instruments for children age 0 to 5 years in US primary care settings. (6)

For broad-band developmental screening tools, from 2002 to 2009, pediatricians’ use of the ASQ increased from 7.3% to 22.4% and use of the PEDS increased from 2.4% to 15.9%. (6) In 2011, a well-designed study indicated that the ASQ (second edition) was significantly more accurate than the PEDS, especially for children over age 30 months, in the identification of developmental delays in a primary care setting. (11) The strengths of the ASQ include its accuracy (both overall and across all age ranges) and its ability to be used as a developmental promotion tool. When caregivers mark “not yet” to an item because they have never tried that particular developmental task with their child, this is considered a “teachable moment,” where clinicians can promote a new developmental activity.

Although multiple primary care studies suggest that the paper-based ASQ is feasible, it does require getting all items on the correct age-interval ASQ thoughtfully completed by caregivers (and then correctly scored) before the health supervision visit. In contrast, the strength of the PEDS is its feasibility as a previst screen with its 2- to 5-minute administration time, fourth- to fifth-grade reading level, and single questionnaire format for all ages.

Interestingly, the PEDS has been designed to detect a broader array of “at-risk” (not just “delayed”) children and can be used not only as a screening and surveillance measure but also as a scaffold to enhance parent–provider communication. Compared with the ASQ, a higher

Figure 3. Mental health/behavior screening & assessment measures (used mostly in school age children). Commonly used ADHD tools were included per the editor’s request; however, an exhaustive list of depression-, anxiety-, and substance abuse-specific screening tools was not included. Core standards are highlighted in yellow. See text to explain shading and see Table 3 for abbreviations.
The strength of the M-CHAT is its feasibility and that it is far less likely to miss a case of autism, compared with a clinician’s unstructured surveillance. (8) Any positive M-CHAT result should lead to an EI/ECSE agency referral and, ideally, an early return office visit for further assessment. Unfortunately, a positive M-CHAT does not necessarily justify an expensive, comprehensive ASD-specific evaluation and EI plan in all cases. Children who screen positive on the M-CHAT plus Follow-Up Interview should be referred for an ASD-specific EI plan and comprehensive evaluation. Although some children will not be diagnosed as having ASD, ultimately over-referrals are not a major concern because those not diagnosed as having ASD generally have other developmental disorders, predominantly language impairment or cognitive disabilities.

Of note, unstructured surveillance (eg, caregiver concerns about ASD; worrisome social-communication deficits or behavioral observations; or having a sibling who has ASD) still should be used in combination with the interpretation of a screening tool (eg, a number of “critical” M-CHAT items failed) to determine the need for an ASD-specific comprehensive evaluation and EI plan without need for the follow-up interview.

To develop a deeper understanding about the strengths and weaknesses of various screening tools, an understanding is needed about the basic precepts behind psychometric and feasibility standards. Core standards affect a screening tool’s application to a primary care setting (eg, performance on a wide variety of populations, screening tool completion rates, sustainability over time). The question arises, what is a core standard? Throughout this article, a core standard is defined as an inflexible property administered routinely for any positive or borderline M-CHAT result due to a high false-positive rate (~90% are not subsequently diagnosed as having ASD) when using the M-CHAT at 18 and 24 months in a general, primary care population. (8)
that is critically important for the effective and feasible performance of a screening measure on a primary care population but is not necessarily dependent on the individual needs of a medical home.

Core standards provide a frame of reference to assure that relevant technical information is provided, with the hope that practitioners thereby will make wiser, more informed decisions when selecting, implementing, and interpreting screening tools. To learn more about what technical properties make a screening tool psychometrically sound, go to http://pedsinreview.aappublications.org/content/33/10/448/suppl/DC1, and to learn what properties make a screening tool feasible, go to http://pedsinreview.aappublications.org/content/33/10/448/suppl/DC2.

Application of Psychometric and Feasibility Standards to Primary Care

The AAP strongly endorses the use of properly standardized, reliable (≥80%), well-validated, and accurate (sensitivity and specificity ≥70%) screening instruments for the early identification of developmental-behavioral problems. However, highly accurate practitioner-administered instruments often possess undesirable feasibility characteristics for busy primary care settings. Practitioners should look hard at how practical any given instrument they are considering will be in the office setting. Suboptimal feasibility lowers screening tool completion rates. Lower completion rates logically hinder early detection rates because early detection is more apt to rely on subjective, unstructured surveillance instead of psychometrically sound screens.

For the early detection of a broad range of developmental delays before kindergarten entrance, parent-report tools such as the ASQ-3, PEDS, and the PEDS:DM are good examples of effective and feasible screening tools that are well-suited to a wide array of primary care settings; however, the use of yes/no checklists such as the Denver Prescreening Developmental Questionnaire II is not recommended. For the early detection of a broad range of social-emotional/mental health problems, the Ages & Stages Questionnaire: Social-Emotional for Growth and Development developmental-behavioral screening measures

Figure 5. Algorithm for developmental-behavioral surveillance and screening (0 to 5 years) in a medical home setting. Permission to print this adapted figure was obtained from Sage Publishing. © Kevin Marks MD, 2012

---

**Figure 5. Algorithm for developmental-behavioral surveillance and screening (0 to 5 years) in a medical home setting.**

**Legend**

- **Screen** = process or action step
- **Diamond** = decision
- **Y** = yes
- **N** = no
- (+) = concerning
- (−) = borderline
- (0) = typical AAP = American Academy of Pediatrics
- ASD = Autism Spectrum Disorder
- CHD = congenital heart disease
- EE = early intervention IDEA = Individuals with Disabilities Education Act
- ODD = oppositional defiant disorder
- RE = Reading Disorder
- RFR = Reach Out & Read
- WCRC = well-child visit

**Before entering the exam room, administer & score psychometrically sound screen(s) per AAP recommendations:**

- **1. Elicit & Address Parents’ & Other Caregivers’ Concerns:**
  - Y = yes
  - N = no

- **2. Gather & Maintain a Dev-Behav. History (milestone/skill monitoring):**
  - Any specific problem pattern that would suggest a disorder?

- **3. Identify Biological & Environmental Dev-Behav. Risk & Protective Factors:**

- **4. Make Accurate & Informed Observations about the Child & Parent-Child Interaction:**

- **5. Promotion of Developmental & Behavioral Wellness:**
  - Literacy (eg, RCR), strength-based, positive parenting/resilience counseling.

- **6. Interpret or Administer/Interpret Screen(s) & Discuss with Parents:**
  - (+) screen
  - (−) screen

- **7. Refer to IDEA (IECSE) & Beneficial Community Programs (eg, Head Start):**
  - If yes for (−), consider referral for multi-specialty dev-behav. evaluation.

- **8. Document Process & Findings & share key findings with referral source(s):**

- **9. Supplemental Care Coordination:**
  - Screen or assessment needed?
  - Return office visit needed?

- **10. Respond to Concerning Screen(s) or Surveillance with Supplemental Screens &/or Medical Tests:**

- **11. Review Referral Feedback Reports & Recommendations while Tracking IDEA (IECSE) Eligibility Status:**

- **12. Monitor Progress & Need for Additional Services after Initiating Chronic Condition Management for Special Health Care Needs:**

---

**AAP Screening Recommendations**

- (a) General developmental screening routinely at 18, 24, or 30 months plus at 4 years to measure “vigilance readiness,” plus as needed if at risk for DD problems.
- (b) Screen routinely at 18 months, plus as needed if at risk for DD problems.
- (c) Autism-specific screening routinely at 18 and 24 months, plus as needed if at risk for ASD.
- (d) Consider family psychiatric screen (eg, GSE 1 mo or at the initial visit), plus routinely employ a maternal mental disorder screen (after first year (eg, 2, 4 and/or 5 mo).

---

**Application of Psychometric and Feasibility Standards to Primary Care**

The AAP strongly endorses the use of properly standardized, reliable (≥80%), well-validated, and accurate (sensitivity and specificity ≥70%) screening instruments for the early identification of developmental-behavioral problems. However, highly accurate practitioner-administered instruments often possess undesirable feasibility characteristics for busy primary care settings. Practitioners should look hard at how practical any given instrument they are considering will be in the office setting. Suboptimal feasibility lowers screening tool completion rates. Lower completion rates logically hinder early detection rates because early detection is more apt to rely on subjective, unstructured surveillance instead of psychometrically sound screens.

For the early detection of a broad range of developmental delays before kindergarten entrance, parent-report tools such as the ASQ-3, PEDS, and the PEDS:DM are good examples of effective and feasible screening tools that are well-suited to a wide array of primary care settings; however, the use of yes/no checklists such as the Denver Prescreening Developmental Questionnaire II is not recommended. For the early detection of a broad range of social-emotional/mental health problems, the Ages & Stages Questionnaire: Social-Emotional for Growth and Development developmental-behavioral screening measures...
For autism-specific screening, the M-CHAT is currently a commonly selected tool. However, whenever the M-CHAT is positive (failed), it is important to administer its follow-up interview to filter out which children are truly in need of an ASD-specific comprehensive evaluation and EI plan.

When selecting instruments, be aware that medical home differences often exist in patient sociodemographic characteristics, patient access, patient flow and volume, office staff resources, state EI/ECSE eligibility criteria, and the availability of early childhood community resources. For all these reasons, different practices may require a different combination of tools to optimize their early detection and referral rates. Table 3 provides a list of commonly used tools, along with their abbreviations and Web links.

### Table 3. Abbreviations and Web Links for Commonly Used Measures

<table>
<thead>
<tr>
<th>Category</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Broad-band developmental: parent-report screening measures</td>
<td></td>
</tr>
<tr>
<td>2. IDI</td>
<td>Infant Development Inventory; <a href="http://www.childdevrev.com/index.html">http://www.childdevrev.com/index.html</a></td>
</tr>
<tr>
<td>3. CDI</td>
<td>Child Development Inventory; <a href="http://www.childdevrev.com/index.html">http://www.childdevrev.com/index.html</a></td>
</tr>
<tr>
<td>4. PEDS</td>
<td>Parents’ Evaluation of Developmental Status; <a href="http://www.pedstest.com">http://www.pedstest.com</a></td>
</tr>
<tr>
<td>5. PDQ-II</td>
<td>Prescreening Developmental Questionnaire II; <a href="http://www.denverii.com/PDQ.html">http://www.denverii.com/PDQ.html</a></td>
</tr>
<tr>
<td>B. Broad-band developmental: practitioner-administered screening or assessment measures</td>
<td></td>
</tr>
<tr>
<td>1. BDIST-II</td>
<td>Battelle Developmental Inventory Screening Test–II; <a href="http://www.riversidepublishing.com">http://www.riversidepublishing.com</a></td>
</tr>
<tr>
<td>2. BITSEA</td>
<td>Brief Infant Toddler Social Emotional Assessment; <a href="http://pearsonassess.com">http://pearsonassess.com</a></td>
</tr>
<tr>
<td>3. Capute Scales;</td>
<td><a href="http://www/curriculumassociates">http://www/curriculumassociates</a></td>
</tr>
<tr>
<td>5. M-CHAT</td>
<td>Modified Checklist for Autism in Toddlers; <a href="http://www2.gsu.edu/~cpps/">http://www2.gsu.edu/~cpps/</a></td>
</tr>
<tr>
<td>C. Narrow-band: social-emotional-, communication-, and autism-specific screening measures</td>
<td></td>
</tr>
<tr>
<td>2. BITSEA</td>
<td>Brief Infant Toddler Social Emotional Assessment; <a href="http://pearsonassess.com">http://pearsonassess.com</a></td>
</tr>
<tr>
<td>D. Mental health screening and assessment measures (for school-age children)</td>
<td></td>
</tr>
<tr>
<td>1. PSC</td>
<td>Pediatric Symptom Checklist; <a href="http://www2.massgeneral.org/allpsych/pscpsc_home.htm">http://www2.massgeneral.org/allpsych/pscpsc_home.htm</a></td>
</tr>
<tr>
<td>2. SDQ</td>
<td>Strengths and Difficulties Questionnaire; <a href="http://www.sdqinfo.com/">http://www.sdqinfo.com/</a></td>
</tr>
<tr>
<td>4. Vanderbilt Assessment Scales; <a href="http://peds.mc.vanderbilt.edu/cdc/childdevelopcenter.htm">http://peds.mc.vanderbilt.edu/cdc/childdevelopcenter.htm</a></td>
<td></td>
</tr>
</tbody>
</table>
using paper-based and, more recently, online screening modalities. Some forms of bias are present with varying degrees of severity within the validation studies for all of these measures, but noteworthy issues are highlighted for each instrument. Under feasibility, costs have been calculated uniformly (the same for each tool) by using a paper-and-pen, modifiable model from Dobrez et al. (13) A cost of $60 per hour was used to account for the combination of practitioner and clinic staff wages. Clinic costs will vary based on implementation procedures, staff salaries, the percentage of caregivers with literacy problems, and the percentage of positive or concerning screens in one’s practice.

Summary

- Although development and behavior are more complex than most anthropomorphic functions that are measured in clinical practice, they are quantifiable, even when applied to a busy primary care setting.
- Informal approaches to eliciting concerns or measuring milestones, such as yes/no checklists extrapolated from lengthier measures, surely contribute to low detection rates of developmental-behavioral problems discerned by primary care providers.
- The periodic use of screening tools that better adhere to core psychometric and feasibility standards (eg, Ages & Stages Questionnaire, Third Edition [ASQ-3], Parents’ Evaluation of Developmental Status [Peds] and/or Parents’ Evaluation of Developmental Status–Developmental Milestones [Peds:DM], Ages & Stages Questionnaire: Social–Emotional, Pediatric Symptom Checklist) helps practitioners to optimize their early identification rates for children who have developmental-behavioral problems, which leads to more effectively and efficiently enrolling greater numbers in early intervention (EI), early childhood special education (ECSE), and other beneficial community services.
- Nevertheless, to achieve a screening tool’s reported reliability and accuracy, it is essential that the tool be administered and interpreted thoughtfully as described in its user’s manual or official Web site.
- Careful attention to proper implementation and interpretation also leads to a more collaborative conversation between parents and practitioners.
- When screening results are concerning, proper implementation and same-day interpretation ensure that referrals occur in a 1) safe, 2) equitable, 3) effective, 4) timely, 5) parent- and patient-centered, and 6) efficient manner, fulfilling the six quality aims of the Institute of Medicine.
- If a practice has failed to implement the periodic use of evidence-based screening measures successfully, then the clinicians should strongly consider selecting some broad-band developmental, social-emotional/behavioral, and autism-specific screening tools for the practice.

Note: To view the references for this article and to view larger versions of the figures, visit the October issue at http://pedsinreview.aappublications.org and click on the “Understanding Developmental-Behavioral Screening Measures” article.

PIR Quiz

This quiz is available online at http://www.pedsinreview.aappublications.org. NOTE: Since January 2012, learners can take Pediatrics in Review quizzes and claim credit online only. No paper answer form will be printed in the journal.

New Minimum Performance Level Requirements

Per the 2010 revision of the American Medical Association (AMA) Physician’s Recognition Award (PRA) and credit system, a minimum performance level must be established on enduring material and journal-based CME activities that are certified for AMA PRA Category 1 Credit™. In order to successfully complete 2012 Pediatrics in Review articles for AMA PRA Category 1 Credit™, learners must demonstrate a minimum performance level of 60% or higher on this assessment, which measures achievement of the educational purpose and/or objectives of this activity.

Starting with the 2012 issues of Pediatrics in Review, AMA PRA Category 1 Credit™ may be claimed only if 60% or more of the questions are answered correctly. If you score less than 60% on the assessment, you will be given additional opportunities to answer questions until an overall 60% or greater score is achieved.

1. A 17-year-old mother has concerns about her 3-year-old boy’s repeated temper tantrums and his frequently ignoring her verbal communication. She is feeling increasingly frustrated. As a primary care provider, which of the following is the most appropriate next step?
   A. Administer a narrow-band ADHD screen.
   B. Administer and interpret a general developmental and social–emotional screen.
C. Refer to a developmental-behavioral specialist.
D. Refer to a teenage mother self-help group for counseling.
E. Schedule a follow-up visit in 6 months and evaluate if symptoms persist.

2. A 2-year-old otherwise healthy girl has not yet started to say “mama” and “dada.” Her motor development appears normal. Mother states that she does not appear to interact normally with her older siblings. Which of the following statements is the most accurate regarding developmental and social-emotional screening in this situation?
   A. False-positive result will unnecessarily increase parental anxiety.
   B. Normal screening results should reassure the parents that intervention will not be needed.
   C. Periodic screening identifies at-risk children for early intervention, but in this clinical scenario, an automatic referral to an EI agency is justifiable.
   D. Screening is not valid in children age <5 years.
   E. Screening should only be performed by a developmental-behavioral specialist.

3. A mother is concerned about her 9-month-old boy developing autism like her neighbor’s child. Which of the following is the most appropriate age to routinely administer an autism-specific screening tool?
   A. 9 months and 15 months.
   B. 18 months and 2 years.
   C. 30 months and 3 years.
   D. 42 months and 4 years.
   E. 54 months and 5 years.

4. Which of the following statements is most correct regarding developmental-behavioral screening measures?
   A. Children who fail screening tests, but subsequently are found not to have a developmental-behavioral problem, later perform as well academically as their peers.
   B. Clinicians’ subjective impressions are more accurate in identifying children having developmental-behavioral problem (sensitivity) than excluding those who do not (specificity).
   C. Majority of developmental delays are correctly identified by clinicians on routine examinations during health supervision visits.
   D. Periodic developmental-behavioral screening has been proven to enhance a clinician’s ability to detect children who have evolving developmental-behavioral problems.
   E. There is no evidence that early identification and intervention of developmental-behavioral problems before kindergarten entrance improves subsequent academic performance.

5. The mother of a 5-year-old boy replies to your inquiry about developmental or behavioral concerns, “Well, he’s a brat sometimes, but I think that’s pretty normal for this age.” To detect a wide spectrum of developmental-behavioral conditions accurately over time, the American Academy of Pediatrics recommends:
   A. Administering a social-emotional screening instrument only as needed if further history suggests that he is at risk for a previously undetected behavioral disorder.
   B. Automatically referring the child and family to a mental health specialist.
   C. Automatically referring the child to assess eligibility for Early Childhood Special Education (ECSE)/Individual with Disability Education Act (IDEA) services.
   D. Routinely administering a screening test for maternal depression.
   E. Routinely administering a social-emotional screening instrument.

HealthyChildren.org Parent Resources From the AAP

The reader is likely to find material to share with parents that is relevant to this article by visiting these links: