# Developmental and Social-Emotional Screening with the Ages & Stages Questionnaire in the United States and Scandinavia: A Systematic Review

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th><em>Developmental Medicine &amp; Child Neurology</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>DMCN-SRE-18-05-0314.R1</td>
</tr>
<tr>
<td><strong>Manuscript Type:</strong></td>
<td>Systematic Review</td>
</tr>
<tr>
<td><strong>Date Submitted by the Author:</strong></td>
<td>n/a</td>
</tr>
</tbody>
</table>
| **Complete List of Authors:** | Marks, Kevin; PeaceHealth Medical Group, Pediatrics  
Madsen Sjö, Nina; National Research Centre for Disadvantaged Children and Youth, University College Copenhagen  
Wilson, Philip; University of Aberdeen, Centre for Rural Health; University of Copenhagen, Centre for Research and Education in General Practice |
| **Keywords:** | developmental screening, social-emotional screening, Ages & Stages Questionnaire, Ages & Stages Questionnaire-Social Emotional, behavioral screening |

Mac Keith Press
Developmental and Social-Emotional Screening with the Ages & Stages Questionnaire in the United States and Scandinavia: A Systematic Review

Authors: Kevin P. Marks MD(1), Nina Madsen Sjö MSc PhD(2), Philip Wilson DPhil FRCPCH FRCGP(3,4)

Affiliations: (1) PeaceHealth Medical Group, Department of Pediatrics, Eugene, Oregon, United States; (2) National Research Centre for Disadvantaged Children and Youth, University College Copenhagen, Copenhagen, Denmark. (3) Centre for Research and Education in General Practice, University of Copenhagen (4) Centre for Rural Health, University of Aberdeen.

Address correspondence to: Kevin P. Marks, MD FAAP kmarksodinson@gmail.com

Short title: Developmental-Behavioral Screening: USA and Scandinavia

Article category: Systematic review.


Key Words: Ages & Stages Questionnaire, ASQ, Ages & Stages Questionnaire:Social-Emotional, ASQ:SE, developmental screening, social-emotional screening, behavioral screening

Funding source: No external funding was secured for this study.

Financial disclosure: The authors have no financial relationships relevant to this article to disclose.

Conflicts of Interest: The authors have no conflicts of interest to disclose.

Main text word count: originally 4998 but then 5379 after revisions. Abstract word count: 200

Authorship: All the authors have read this manuscript and agreed to its being submitted for publication. All individuals listed as authors meet the appropriate authorship criteria and nobody who qualifies for authorship has been omitted from the list. Dr. Marks conceptualized this systematic review, helped to perform the systematic review, wrote and revised the manuscript. Dr. Madsen Sjö helped to perform the systematic review and reviewed and revised the manuscript. Prof. Phil Wilson reviewed and revised this manuscript. Each author had complete access to the study data that support this study.
ABSTRACT

Aims: To investigate screening practices with the Ages & Stages Questionnaire (ASQ) and ASQ:Social-Emotional (ASQ:SE) in the USA and Scandinavia and to identify practical lessons and research opportunities. Methods: A systematic review was performed about ASQ/ASQ:SE-related studies in children 0 through 5 years. From 9 databases and 1689 references (published 1988 to 2018), 127 articles were included and categorized using Covidence online software. The Critical Appraisal Skills Programme Checklists were used before data-synthesis. Results: US studies primarily use the ASQ/ASQ:SE to detect delays in general and at-risk populations in medical settings, which increases early detection, clinician-referral and intervention rates. Scandinavian studies commonly use the ASQ/ASQ:SE to monitor developmental-behavioral differences in intervention/exposure-based cohorts. Pre-visit screening yields completion/return rates of 83% to over 90% and fosters same-day interpretation. When referrals are indicated, systemwide care coordination or coloclation with a developmental-behavioral specialist is beneficial. Research opportunities include investigating/measuring the ASQ/ASQ:SE’s “overall” sections. Danish, Norwegian and Swedish translations are available but up-to-date norming and validation studies are needed throughout Scandinavia. Randomized controlled trials are needed to investigate outcomes in screened versus unscreened cohorts. Interpretation: Practical implementation lessons are reviewed. The ASQ/ASQ:SE authors should consider making the “overall” sections measurable. Many research opportunities exist.

What this paper adds

- General and at-risk populations broadly benefited from periodic ASQ and/or ASQ:SE screening.
- Pre-visit (not mail-back, post-visit) ASQ and/or ASQ:SE implementation systems work best.
- The ASQ and ASQ:SE “overall” sections are not quantifiable and under researched.
- Up-to-date ASQ and ASQ:SE psychometric studies are needed across Scandinavia.
- ASQ/ASQ:SE randomized controlled trials are needed with developmental-behavioral, health and socio-economic outcomes.
Introduction

This systematic review’s primary research question is: 1) How do different developed countries (the United States, Denmark, Norway and Sweden) use the same two parent-centered, developmental-behavioral screening tools, the Ages & Stages Questionnaire (ASQ) and ASQ:Social-Emotional (ASQ:SE) in children from birth through five years? These developed Western nations lie farther apart on the spectrum of preventative care models. Findings might be generalizable to other developed Western countries that possess a mixture of the US and Scandinavian models of preventive care. A secondary question is: 2) What is known about the effectiveness, feasibility, implementation, considerations for at-risk populations, follow-up steps and research gaps of universal ASQ/ASQ:SE screening?

An estimated 12% of US children are diagnosed with any developmental disability between 3 to 10 years, and 16% of children between 11 to 17 years.\(^1\) An estimated 21% of US children aged 9 to 17 years have a mental health disorder according to national parent-report surveys.\(^2\) In Scandinavia (specifically Denmark), 12% of 8 to 9-year old children have a developmental disability and/or mental health disorder according to diagnostic criteria.\(^3\) The systematic early identification of children who need early intervention (EI) services can benefit those with emerging developmental delays, emerging social-emotional/behavioral problems, and a history of exposure to adverse childhood experiences.\(^4,5\) Evidence-based EI can produce individual, family and societal benefits.\(^4-6\)

In the United States, economic evaluations have indicated high returns on investment for birth-through-five developmental programs, and the younger the high-quality interventions are provided, the higher the return rate.\(^6\) This is one of the many reasons that the American
Academy of Pediatrics (AAP) recommends universally screening for maternal depression/anxiety at 1, 2, 4 and 6 months, developmental delays at 9, 18 and 24-30 months, autism at 18 and 24 months, and social-emotional/behavioral problems when concerns arise from birth to five years.\(^{7-10}\) In reality, in 2017/2018, only 30.4% of U.S. children received a parent-report, standardized developmental screening and state-level variance in screening rates spanned 40 percentage points (17.2% in Mississippi and 58.8% in Oregon).\(^{11}\) Primary care providers (pediatricians, family practitioners) must also rely upon their surveillance to assess risk level to determine whether a standardized screening is necessary at any well-visit where one of the above standardized screenings is not routinely recommended.\(^{8-10}\)

Whenever at-risk children are identified, they can be referred to a diverse assortment of services, including Individuals with Disabilities Education Act (IDEA) or early intervention/early childhood special education (EI/ECSE) agencies, parenting programs, daycare provisions, early learning/preschool, developmental-behavioral pediatricians, mental health providers, and other professionals. A small minority of moderate to high-risk infants and toddlers are seen by home visit programs. About half of children aged 3 to 5 years are enrolled in preschool, which is not universally subsidized by the US government, resulting in enrollment disparities based on socio-economic and disability status.\(^{12}\) Heterogeneity in policies and services, along with “mismeasurement, misuse, and mismanagement” problems have been said to hamper “systemwide solutions” across the US early detection and intervention system.\(^{13,14}\)

Scandinavia (here defined as Denmark, Norway and Sweden) has a different model of preventive care. All families are provided paid parental leave in the first year of life, subsidized daycare and affordable early learning/preschool. Country-specific early detection policies should be viewed in this context.\(^{15}\) Centralized authorities do not currently recommend universal,
broad-band developmental-behavioral screening using psychometrically sound, parent-centered instruments.\(^{(15)}\) Instead, practitioner-administered screening is variably performed by nurses or doctors using a variety of narrow-band instruments (e.g., the Alarm Distress Baby Scale, which measures social withdrawal behavior, or specific tests of motor function). Language-specific screening is typically performed at 2.5 to 3 years of age in preschools, child health centers or in general medical practices. Children are universally and periodically seen by nurses and/or doctors in the first six years of life at home, in child health centers or general medical practices according to local policies. The content of scheduled health maintenance visits follows country-specific policies regarding “developmental surveillance” and hands-on “neurodevelopmental assessments” which may be non-standardized and without clear cutoffs indicating a need for further evaluation. Medically-focused developmental assessments are generally performed by general practitioners in their practices or in child development centers. Children only see medical subspecialists (which includes pediatricians) on an as needed basis. Ninety-three to 96% of children aged 3 to 5 years are enrolled in preschool throughout Scandinavia.\(^{(12)}\) Nearly universal preschool attendance may help to improve the early detection of developmental-behavioral problems. Equitable, affordable early childhood education and other prevention-based services increase social and income mobility,\(^{(16)}\) which might decrease the prevalence (and severity) of developmental-behavioral problems.

However, before comparing how different countries and cultures use screening tools in various settings, it is first important to review concepts and evidence regarding developmental-behavioral surveillance/monitoring and screening. To best identify which children need EI/ECSE, surveillance should always complement screening and vice versa.\(^{(17,18)}\) Surveillance refers to a flexible, longitudinal, continuous and cumulative process whereby knowledgeable
professionals have multiple action steps and decision-making points
(http://archive.brookespublishing.com/documents/Bricker-screening-algorithm.pdf).\(^{17,19,20}\)

Developmental-behavioral screening is a “brief, formal evaluation of developmental-behavioral
skills intended to identify those children with potential problems who should be referred for
further assessment”.\(^{19,20}\) The number of pediatric screens that meet this definition in the US and
Scandinavia is relatively large. Different screens typically measure one of the first four
components of surveillance: 1) identifying and addressing parents’ concerns, 2) gathering and
maintaining developmental-behavioral milestones/skills, 3) identifying environmental and
biological risk and protective factors, 4) making informed observations about the child and
parent-child interaction.\(^{7,17,19,20}\) Single-point screening (e.g., at 2.5 years only) might identify a
suspected developmental delay (the acquisition of milestones/skills at a significantly slower-
than-normal rate) more accurately but not problem patterns over time. Planned, periodic
screening is recommended because, in addition to identifying or monitoring delays, medical
professionals can better identify a developmental-behavioral dissociation (when milestones/skills
are occurring in one area but not another), a deviancy (when milestones/skills occur out of
sequence) or a regression (when there is a halted acquisition or plateau of milestones/skills),
which can raise the index of suspicion for an iron deficiency, autism spectrum disorder, motor
disorder, sleep disorder, sensory impairment, genetic or metabolic disorder, epilepsy, child
neglect and/or abuse, among other conditions which might require intervention. In the US,
Scandinavia and other developed Western nations, it is the responsibility of medical
professionals to identify and address developmental-behavioral disorders and co-occurring
medical conditions.
For this systematic review, the ASQ and ASQ:SE have been selected to compare screening practices across the U.S. and Scandinavia. After scanning of the literature, it was discovered that they are the only two screens for infants, toddlers and preschoolers that have been translated and applied in research or clinical use in the U.S., Denmark, Norway and Sweden. Analyzing the use of these two screens allows for a more consistent comparison of developed nations interested in the early detection of a broad spectrum of developmental-behavioral (often coexisting) conditions.

The ASQ-3 is a 30-item, parent-report screen that accesses five domains of development (communication, gross motor, fine motor, problem solving, personal-social) and gathers 6 to 9 overall concerns for children 1 through 66 months of age.\(^{(21)}\) It has sensitivity=82.5%-89.2% and specificity=77.9%-92.1% (mostly using the Battelle Developmental Inventory, 2nd edition\(^{(22)}\) as the criterion-based instrument), interobserver reliability=0.93, test-retest reliability=0.75-0.82, internal consistency=0.6-0.85 and was normed on 12,695 children (a range of 352 to 2088 children per age-interval, aged 2 months through 5 years) representative of the US population.\(^{(21)}\) The ASQ-3’s concurrent validity has been closely replicated in a primary care medical setting\(^{(23)}\).

The ASQ:Social-Emotional-2 (ASQ:SE-2) is a 19- to 33-item, parent-report screen that accesses seven social-emotional areas (self-regulation, compliance, communication, adaptive functioning, autonomy, affect, interaction with people) and has clear typical/monitor/refer cutoffs for children 3 through 65 months.\(^{(24)}\) It has sensitivity=70.8%-84.6% and specificity=89.5%-98.2% (compared to three criterion-based instruments and/or a professional diagnosis of a social-emotional disability or autism spectrum disorder), test-retest reliability=0.89, internal consistency=0.84, and was normed on 16,394 children (a range of 148
to 1,456 children per age-interval, aged 2 months through 5 years) who were representative of
the US population.\textsuperscript{(24)}

The ASQ and ASQ:SE use parent-report items which allow for “teachable moments” where
the screening process encourages parents to spend “special time” playing with their child if they
have never tried a developmental task.\textsuperscript{(19-21,24)} Compared to other early childhood screeners, the
ASQ and ASQ:SE have a wide age range, birth-through-5-years. They are feasible for use in
multiple settings and take 10 to 20 minutes for parents/caregivers to complete.\textsuperscript{(19-21,24)} A previous
systematic review about the ASQ focused on its psychometric properties at 2 to 2.5 years of
age.\textsuperscript{(25)} A narrative review focused on its use in low and middle-income countries.\textsuperscript{(26)} This is the
first systematic review comparing the use of the ASQ and ASQ:SE in the U.S. and other
developed Western countries.

\textbf{Methods}

\textbf{Inclusion/exclusion and categorization criteria}: Refer to Table 1. In this systematic review,
three versions of the ASQ (ASQ, ASQ-2 and ASQ-3) and two versions of the ASQ:SE (ASQ:SE
and ASQ:SE-2) were included, as well as five language versions (English and Spanish for US
studies and Danish, Swedish and Norwegian for Scandinavian studies). The terms ASQ and
ASQ:SE will hereafter be used for general matters that cover all published versions of the two
questionnaires as well as the different language versions. These evolving versions of the ASQ
and ASQ:SE’s item content, result categories and psychometric properties have gradually
changed over time as its authors have re-normed, re-validated and improved their screening tools
every decade. For further details about the evolving differences and how the ASQ and ASQ:SE
can be used together, refer to their manuals and https://agesandstages.com.\textsuperscript{(21,24)}
**Search strategy**

Literature searches were conducted in six international databases (Medline/Pubmed, PSYCHInfo, Embase, Cochrane, Cinahl and ERIC) and three Scandinavian-specific databases (Christin, Forskningsdatabasesn and Diva) published from 1988 to December 21st, 2017. Search terms were: "Ages and Stages Questionnaire" OR "Age and Stage Questionnaire" OR "Ages and Stages Questionnaires" OR "Age and Stage Questionnaires" OR "Ages & Stages Questionnaire" OR "Age & Stage Questionnaire" OR "Ages & Stages Questionnaires" OR "Age & Stage Questionnaires" OR “ASQ*”. As many questionnaires have the abbreviation, ASQ, e.g., Adolescent Stress Questionnaire, the search terms in the largest international databases were restricted to studies in the four target countries and to humans below age 18. To supplement the databases, grey literature was consulted (www.agesandstages.com), as well as presentations at two international conferences solely discussing ASQ and ASQ:SE research. Finally, the lead author of the ASQ and ASQ:SE, Dr. Jane Squires, was contacted and her comprehensive list of published ASQ and ASQ:SE articles was cross-checked against those already identified.

**Data screening, extraction, categorization:** Refer to Figure 1 (PRISMA flow diagram). All identified references were imported into Covidence systematic review online software, Veritas Health Innovation, Melbourne, Australia at https://www.covidence.org. Working independently and in parallel, the first and second authors screened all titles and abstracts, reviewed all full text articles, resolved inclusion/exclusion discrepancies, and categorized reasons for exclusion. For Table 2 (US studies) and Table 3 (Scandinavian studies), extracted data from the included and appraised articles focuses on the original and primary purpose of the ASQ and ASQ:SE, which is universal developmental-behavioral screening in general and at-risk populations.
Critical appraisal and data-synthesis:

The Critical Appraisal Skills Programme (CASP) Checklists [https://casp-uk.net](https://casp-uk.net) were used to assess the risk of bias in eligible, targeted studies.\(^{(27)}\) See Tables 2 and 3. When doing data-synthesis, the more credible or less biased studies were favored before forging conclusions. A meta-analysis was not completed because: 1) our objectives were broader clinical questions. 2) Studies with reference/gold/diagnostic standard evaluations were relatively sparse and typically only investigated outcomes on the children with problematic screening results. 3) Outcomes were qualitatively different so data could not be easily combined to increase sample size and estimate effect size.

**Results**

**Overall findings and research opportunities:** See Figure 1 (PRISMA). Studies describing the impact of universal screening are presented in Tables 2 and 3. No studies investigated harms or reported the ASQ and/or ASQ:SE causing harm. Many excluded studies did not adhere to the directions in the ASQ and ASQ:SE User’s Manuals. Eleven studies did not use all five ASQ subdomains. Four studies used abbreviated versions of the ASQ subdomains or an abbreviated version of the scored portion of the ASQ:SE. No articles measured and analyzed the results of the ASQ and/or ASQ:SE’s “overall" section. Because the overall section cannot be scored, the first and second authors decided not to exclude all studies where the overall questions were not taken into account.

Figures 2 and 3 quantify how the ASQ and ASQ:SE are being used/studied in Scandinavia and the US in a variety of different settings. US studies \((n=90)\) are more commonly using the ASQ and/or ASQ:SE for the early detection of developmental-behavioral delays in general \((n=32)\) or at-risk \((n=20)\) populations—most commonly in primary care medical settings
Scandinavia had 20 intervention/exposure-based studies that longitudinally tracked developmental-behavioral differences (using the ASQ and/or ASQ:SE) between intervention/exposure and control groups. The US had only 12 studies in the intervention/exposure research category. Tables 2 and 3 focus on articles that employed the ASQ and ASQ:SE for its original purpose, i.e., universal developmental-behavioral screening in general or at-risk populations.

**Effectiveness and feasibility in US settings:** North Carolina’s (Assuring Better Child health and Development or ABCD) initiative steadily increased the number of ASQ-2 screenings per child over the first 2 to 2.5 years of the project, which led to children being referred to EI at earlier ages. The percentage of children receiving EI statewide increased from 3.0% to 4.3%.

An Illinois statewide project trained primary care practices to routinely and periodically screen with the ASQ-2 and ASQ:SE in the first three years of life. The project reached its target of screening 85% of patients with the ASQ-2 at 11 out of 16 participating practices in children 1 and 2 years of age. At baseline, only 4 out of 16 clinics had routinely done developmental screening in 1-year-olds, and 2 out of 16 clinics in 2-year olds. ASQ:SE screening rates at 18 months increased to 85% in 7 out of 16 of participating practices. At baseline, only 1 out of 16 clinics was routinely doing social-emotional screening. These two statewide initiatives, along with two other studies, found the overwhelming majority of parents/caregivers and physicians felt screening was feasible, covered important areas of child development and uncovered new information about children’s strengths and limitations.

Seven studies in general populations found the periodic use of the ASQ and/or ASQ:SE is feasible, low cost and increases early identification and EI/ECSE eligibility rates. Pediatricians (without the ASQ and/or ASQ:SE) either habitually miss or take a “wait and see”
approach with children who possess suspected developmental delays and/or social-emotional problems who would otherwise qualify for EI/ECSE services. The Hix-Small 2006 study found the ASQ-2 led to a 5.7-fold increase in EI referrals at 12 months and a 2.9-fold increase at 24 months, compared to a control year, when pediatricians were relying on their unstructured surveillance alone.\(^{(42,43)}\) Periodic ASQ-2 screening also increased EI eligibility rates.\(^{(28,42,43)}\) The Briggs 2012 study found periodic ASQ:SE screening in an urban primary care setting to be feasible, sustainable over a 5-year period, and effective in addressing concerning ASQ:SE screening results.\(^{(39)}\) However, effectiveness was conditional upon a model of care which included the colocation of early childhood psychologists who provided brief interventions inside the medical practice.\(^{(32,39)}\) The Guevara 2013 randomized controlled parallel-trial (on a general population) found the ASQ-2 significantly decreased the timeline for primary care providers to detect and refer children with suspected developmental delays.\(^{(44)}\) ASQ-2 screenings at 9, 18 and 24-30 months significantly shortened the timeline to linking at-risk children/families to EI agencies so they received services at younger ages.\(^{(44)}\)

**Implementation lessons in US settings:** Physicians/practices improved office-based ASQ-2 implementation using a team-based approach where schedulers, receptionists, medical assistants and other staff each had assigned tasks so primary care physicians could review scored results prior to entering exam rooms.\(^{(28)}\) Although the ASQ-3 and ASQ:SE-2 User’s Manuals\(^{(21,24)}\) discuss the option of a mail-back protocol, a general population study found a 54% ASQ completion/return rate and called for alternative implementation strategy.\(^{(42,43)}\) North Carolina and Illinois statewide initiatives, along with a medical home for foster children in New York state, used office flow procedures that expected caregivers to complete the ASQ or ASQ:SE before visits. A pre-visit implementation strategy yielded completion/return rates between 83%
to over 90%. Pre-visit screening, immediate scoring, and face-to-face interpretation of the ASQ and ASQ:SE offer a more effective implementation strategy.

Nevertheless, one study (in a lower risk, general population) found that 12% of parents/caregivers found pre-visit online screening (CHADIS https://www.chadis.com/site/) to be “somewhat or very difficult to use.” Another study found the paper–pencil versus online ASQ results to be considered roughly equivalent or interchangeable. Two randomized trials found a measurable, but not statistically significant difference between the ASQ being completed by parents in the waiting room without assistance versus under standardized conditions with trained office staff and ASQ toys.

**Considerations for at-risk populations in US settings:** ASQ and ASQ:SE screening enhances surveillance on at-risk populations, most especially children with adverse childhood experiences including maltreatment or having a deployed military parent. For example, in a specialized medical practice for children “new to foster care”, systematic ASQ-2 screening doubled the detection rates of developmental delays compared to a clinician’s surveillance alone (37% vs. 14% in infants; 89% vs. 42% in toddlers; 82% vs. 44% in preschoolers; 58% vs. 29% overall; \( P \leq .001 \) for all age groups). In this same practice, periodically using the ASQ:SE more promptly detected 6 times more children (ages 6 months to 5.5 years) with suspected social-emotional problems compared to a clinician’s surveillance alone. Use of the ASQ:SE detected 24% of children as having a suspected social-emotional problem, while a clinician’s surveillance detected 4%. The ASQ:SE detected significantly more children with social-emotional problems than the ASQ-2 in this high-risk sample. However, these two studies lacked reference/gold standard assessments. One study found
the ASQ may not adequately identify additional developmental delays (beyond a communication
delay) in young children already diagnosed with bilateral sensorineural hearing loss.\(^{(63)}\)

**Follow-up steps in US settings:** When clinician-based referrals are indicated, systemwide care
coordination or colocation with a developmental-behavioral specialist is beneficial.\(^{(28,33,42,43,64,65)}\)

For example, in a study using the ASQ at 12 and 24 months, two to three years after the initial
EI/ECSE referral, 57% were not deemed eligible for EI/ECSE services.\(^{(42,43)}\) This 57% statistic could make readers mistakenly assume there was an over-referral problem. However, within this 57% group, 13% were placed on a “monitoring list”, 22% were
“screened out” (primarily by phone interviews, without receiving an evaluation), 7% had
“no parental concerns” and/or refused services, and 15% were lost to follow-up.\(^{(42,43)}\) The
true percentage of EI/ECSE over-referred children was unclear, especially since the study was
conducted in Oregon, which has strict IDEA EI/ECSE eligibility criteria, meaning that children
had to have a more pronounced developmental delay (compared to other US states) in order to
qualify for EI/ECSE services.\(^{(42,43)}\) Referral and follow-up problems were again highlighted in
the 2010 King study.\(^{(64)}\) Pediatricians from 15 different states made EI referrals 72% of the time
when ASQ-2 results were abnormal.\(^{(64)}\) Seventeen diverse practices struggled to track EI
referrals and link red-flagged children to EI agencies.\(^{(64)}\)

The Hardy 2015 study investigated the ASQ-3’s pattern profile to determine which
children need a supplemental screening and/or more in-depth follow-up assessment.\(^{(65)}\) Although
further research is needed, the ASQ-3 could possibly be used as a first-level screener to
determine, if an autism-specific screen like the Modified Checklist for Autism in Toddlers-
Revised with Follow up (MCHAT-R/F) is needed.\(^{(65)}\) If the ASQ-3 is typical in all domains, a
supplemental MCHAT-R/F screening is likely not necessary at 18 and 24 months.\textsuperscript{(64)} However, if the ASQ-3’s communication (and/or personal-social) domain scores fall into the referral or monitoring zones, then the MCHAT-R/F should be completed.\textsuperscript{(65)}

**Effectiveness, feasibility and research opportunities in Scandinavia:** Studies generally support the ASQ/ASQ:SE’s ability to effectively and feasibly detect children with suspected developmental-behavioral problems\textsuperscript{(66,67)} and determine typical developmental pathways.\textsuperscript{(68)} However, no Scandinavian ASQ-3 norming and/or validation studies, or ASQ:SE-2 norming studies were found in our literature search. Norway has five norming and/or validation studies which used the older ASQ-2.\textsuperscript{(66,69-72)} One Danish translation ASQ:SE-2 psychometric study,\textsuperscript{(73)} and two Danish ASQ-2 psychometric studies\textsuperscript{(74,75)} were identified. One Swedish ASQ:SE validation study\textsuperscript{(76)} was identified. The Norwegian 6-month ASQ-2 and Danish ASQ:SE-2 were found to have mean scores and cut-offs that are slightly lower compared to US norms.\textsuperscript{(69,73)} Other studies demonstrate that, across age-intervals, the Norwegian ASQ-2 had mean domain scores and cut-offs that were fairly comparable to US normative data, although the domain score variation was smaller in Norwegian samples.\textsuperscript{(70,77)} The 2004 Janson ASQ-2 norming study utilized a randomized sample from Norwegian population,\textsuperscript{(71)} however, it is out-of-date. The Valla 2015 study supported the ASQ-2’s construct validity but recommended gender dependent norms\textsuperscript{(67)} to improve effectiveness; however, the English ASQ-3 was published in 2009.\textsuperscript{(21)}

**Implementation lessons in Scandinavia:** The Janson 2003 study, which used a randomized sample from the Norwegian population, found the ASQ mail-back collection strategy to be problematic as children grew older.\textsuperscript{(77)} ASQ completion rates were higher (76%) in young infants and toddlers but unacceptably low (32%) in preschoolers.\textsuperscript{(77)}

**Considerations for at risk populations in Scandinavia:** The Alvik 2014 study suggested
routinely screening with the 6-month ASQ if the infant has specific risk factors: higher maternal age, having older siblings, and a history of maternal depression.\(^{(78)}\) The Junge 2017 study used the Edinburgh Postnatal Depression Scale to screen mothers for peripartum depression and, when results were positive/concerning, an ASQ:SE screening was recommended for children under two years old.\(^{(79)}\)

**Follow-up steps in Scandinavia** for concerning ASQ/ASQ:SE results were quite inconsistent, compared to US studies where, at a minimum, these children were routinely referred to IDEA EI/ECSE programs (which are available in every county of every state). Scandinavian children with concerning screens often had no clearly described follow-up plan or they received further assessment (typically with a psychologist or medical sub-specialist), based on individualized research protocols.\(^{(66-69,71-99)}\) The literature also collectively indicates different follow-up procedures in different regional municipalities throughout Denmark, Norway and Sweden.\(^{(66-69,71-99)}\) To the best of our knowledge, no recent Scandinavian articles describe national prevalence rates (or source) of early referral to interventions for children with suspected developmental and/or social-emotional delays.

**Discussion**

A main finding of this systematic review was that US professionals more commonly use the ASQ and/or ASQ:SE as it was originally intended—i.e., the early detection and referral of developmental-behavioral concerns in general or at-risk populations, most commonly in primary care medical settings. Scandinavian cohort studies more commonly report the use of the ASQ and ASQ:SE to longitudinally track differences between intervention/exposure groups compared to control groups.\(^{(79-99)}\) There are few studies of these tools being used routinely in clinical
practice across Scandinavian countries, which lack up-to-date, national norms for the ASQ-3 and
ASQ:SE-2.

While 28% (36/127) of all included studies used the ASQ and/or ASQ:SE to universally
screen a general population, 19% (24/127) used them to universally screen an at-risk population.
For all general and nearly all (14/15) at-risk populations, the ASQ and/or ASQ:SE was found to
be beneficial in some way. No studies specifically investigated or reported that screening harmed
children, parents/caregivers or society. Although estimated ASQ-3 costs in the US range from
$2.36 to $14.36 per screening and ASQ:SE-2 costs range from $2.40 to $12.40 per screening,
parent-report screening is far less costly than practitioner-administered screening.\(^{(20,34,100)}\) The
literature also reports that doctors and parents/caregivers value and learn helpful new information
from the screening process.\(^{(28-32)}\)

Precautionary lessons were learned from ninety-one excluded articles. Many studies
“mismeasured” developmental-behavioral skills because they “misused”\(^{(13)}\) the ASQ or ASQ:SE
by not adhering to the directions in their respective User’s Manuals.\(^{(21,24)}\) Use of haphazardly
modified or abbreviated versions of the ASQ and/or ASQ:SE should be discouraged because
concurrent validity, reliability and other psychometrics were calculated using the entire scored
portions of the ASQ and ASQ:SE.\(^{(21,24)}\)

Surprisingly, no articles quantified or analyzed the results of the ASQ and/or
ASQ:SE’s “overall” section when interpreting their data set. These 6 to 9 yes/no questions,
which provide an opportunity for parents/caregivers to write comments about the quality of a
child's development-behavioral skills or risk factors, are routinely being ignored by researchers.
Research is needed to see if the overall section is also being ignored by clinicians and other
professionals before interpreting screening results. The ASQ-3 and ASQ:SE-2 User’s Manuals (21,24) recommend that overall concerns call for follow up. Therefore, the authors of the ASQ and ASQ:SE should consider improving their tools by making the “overall” section quantifiable or measureable. This would allow researchers to investigate how these questions affect its validity.

US studies collectively found that the ASQ and ASQ:SE 1) are feasible across the medical, educational and social service sectors, 2) increase the early detection and referral of suspected developmental-behavioral problems, and 3) increase EI/ECSE eligibility rates in general (21,23,24,28-39,42-46) and higher risk (40,41,47,55-61) populations, especially children with adverse childhood experiences. (40,41,47,55-59,61) Unsurprisingly, the most populous country with the highest percentage of children with developmental-behavioral problems (and where the ASQ and ASQ:SE were developed), had the highest number of studies in every article category. This finding is likely related to an over two-fold increase in the number of US pediatricians using the ASQ between 2002 to 2009. (101) Multiple AAP recommendations, (7-10) nationwide initiatives (e.g. Birth to Five Watch Me Thrive! (14) https://www.acf.hhs.gov/ecd/child-health-development/watch-me-thrive, Healthy Steps (25) https://www.healthysteps.org, Nurse-Family Partnership https://www.nursefamilypartnership.org) and expanding statewide initiatives (e.g. Assuring Better Child Development (28) https://nashp.org/abcd-history/ or Help Me Grow https://helpmegrownational.org) have fueled ASQ and ASQ:SE-related research.

While Scandinavia has published an expanding number of intervention/exposure-based studies (79-99) over the last decade, up-to-date psychometric studies would help to better support the validity of their findings. Norway has provided the most supportive psychometric
Denmark is in the middle. Sweden is lagging behind. The literature suggests the ASQ and ASQ:SE are less commonly used in Scandinavian primary care medical settings, possibly because the Danish, Norwegian and Swedish ASQ and ASQ:SE lack up-to-date norms while research projects are more feasible using screens without up-to-date norms. Nevertheless, it is unclear if the Scandinavian referral/monitor/typical cutoff scores for the ASQ-3 and ASQ:SE-2 should be similar to US cutoff scores across all age-intervals.

If Scandinavia moves forward with universal ASQ/ASQ:SE screening in clinical settings, pre-visit (preferably online) screening is the best strategy. Scandinavian and US studies independently agree the ASQ mail-back collection strategy can be problematic, especially in preschoolers. The Janson 2003 study noted an ASQ completion rate of 76% in infants and toddlers and only 32% in preschoolers. These findings mirror North Carolina’s ABCD project—acceptable ASQ completion rates in infants and toddlers (76%) but unacceptable completion rates in preschoolers (38%).

Evidence from the included articles, combined with knowledge of the Scandinavian model of preventive care, highlights many opportunities. First steps include norming the latest versions of the ASQ and ASQ:SE and then evaluating their impact in home visit, primary care medical and early learning/preschool settings. Unlike the US, Scandinavian general practitioners less commonly use the ASQ and/or ASQ:SE to determine which children need a more in-depth developmental-behavioral assessment or to assist with medical decision-making. The barriers to implementing parent-centered, broad-band developmental-behavioral screening in medical settings is fodder for future research.
Experts purport a “Systems Approach” to universal ASQ/ASQ:SE screening. This means that nationwide initiatives should 1) define program goals, 2) increase community awareness about the benefits of early detection and intervention, 3) provide centralized referral contacts, 4) train practitioners to perform developmental-behavioral screening in combination with activities that promote healthy child development and behavior, 5) leverage existing resources across sectors, and 6) map out the most effective follow up steps for children with suspected delays and at-risk conditions. Given the literature strongly suggests a poorly organized follow-up structure for children with suspected problems, our hypothesis is that a “Systems Approach” ([http://archive.brookespublishing.com/documents/developmental-screening.pdf](http://archive.brookespublishing.com/documents/developmental-screening.pdf)) or the “Help Me Grow System Model” ([https://helpmegrownational.org/what-is-help-me-grow/hmg-system-model/](https://helpmegrownational.org/what-is-help-me-grow/hmg-system-model/)) is lacking but might be beneficial in Scandinavia, the entire US and most likely, other developed Western nations.

**Limitations** include: 1) The scope of this review was restricted to two screening instruments, the ASQ and ASQ:SE. We may have seen a different pattern in screening procedures across four developed countries if different screens were studied because differently designed developmental-behavioral screens can capture significantly different groups of children. 2) Neatly categorizing full text articles into various article types (see Table 1) was sometimes difficult. In instances where categorization was hazy, the authors selected the category closest fit to the article’s objectives. 3) In three out of nine databases, we applied search terms that specified that studies had to be conducted in one of the four target countries as well as applied to children below age 18 years. In case these search terms were not up-to-date in the databases, we may have overlooked studies. On the other hand, to the best of our knowledge, our review has the most advanced combination of search words for ‘Ages and Stages Questionnaire’ compared
to other ASQ/ASQ:SE reviews. 4) Our review was limited to the US, Denmark, Norway and Sweden and caution should be emphasized with interpretation of the smaller body of Scandinavian studies. Nevertheless, it is possible that our findings are generalizable to other developed, higher income Western nations.

Conclusions

What is already known but has been reaffirmed?

1. Universal and periodic ASQ and/or ASQ:SE screening increases the early detection and referral of suspected developmental-behavioral problems, and EI/ECSE eligibility rates\(^{(28,30,39-43)}\) in general\(^{(23,28-32,39,42-44)}\) and higher risk\(^{(40,41,47-62)}\) populations. One randomized controlled parallel-trial found that ASQ screening at 9, 18 and 24-30 months in a primary care setting significantly improves the timeline to early detection, referral and linkage to EI services.\(^{(44)}\)

2. Periodic ASQ and ASQ:SE screening is feasible and sustainable in primary care medical settings.\(^{(28-30,39-44)}\)

3. All screening programs have economic costs and can potentially cause harm\(^{(149)}\) and therefore, implementation procedures should promote healthy developmental activities in hope of lessening the potential for causing harm. Although publication bias might be an explanation,\(^{(149)}\) no studies explicitly investigated harms or reported the ASQ and/or ASQ:SE harmed children, parents/caregivers or society. Five studies reported that medical clinicians and parents/caregivers value and learn important new information (about children’s strengths and limitations) from the screening process.\(^{(28-32)}\)
4. A measurable, but not statistically significant difference was not found between ASQ results when the ASQ was completed in the waiting room without assistance versus under standardized conditions with ASQ-3 toys and trained office staff, according to two randomized controlled trials.\(^{(44,46)}\)

5. One well-conducted study found that the colocation of early childhood developmental-behavioral specialists inside medical practices improves short-term outcomes and more effectively fosters referral and care coordination services for children with positive/concerning ASQ:SE results.\(^{(39)}\)

6. When a EI/ECSE or community-based referral is indicated, systemwide care coordination is needed to ensure that further evaluation reliably occurs in a timely manner and the most effective services are provided.\(^{(14,17,19,33,64)}\) System-wide care coordination, a clinician’s surveillance and supplemental screening can hone referral choices and provide ongoing monitoring.\(^{(17,19,30,65)}\)

7. Family-centered, high-quality EI/ECSE can potentially improves outcomes.\(^{(4-6,19)}\) Periodic screening increases early detection and the percentage of children receiving EI/ECSE services.\(^{(28,30,39-43)}\) However, no studies have “connected the dots” by proving that periodic ASQ and/or ASQ:SE screening improves developmental-behavioral, health and socio-economic outcomes compared to non-screened populations. Longitudinal, prospective, randomized control trials are needed in Scandinavia and the US examining the effects of universal screening.

*What is new?*
1. Studies that used the mail-back protocol had suboptimal ASQ/ASQ:SE completion/return rates ranging from 32% to 76% (42,43,77) An pre-visit screening implementation system, with immediate scoring and provider-to-parent interpretation, yields completion/return rates of 83% to over 90% (28-30,40,41).

2. Fourteen higher risk groups of children benefited from periodic ASQ and/or ASQ:SE screening. These groups included: exposure to peripartum maternal depression, foster care placement, parents not being proficient in country’s primary language, homelessness, teen parents, international adoptees, children frequently presenting to emergency rooms, children from military families with a deployed parent, preterm or low birth weight, epilepsy/seizure disorder, complex congenital heart disease, sickle cell disease, cancer, and a positive newborn metabolic screening result that could adversely affect neurodevelopment (40,41,43,47-62). One outlier study found the ASQ did not effectively identify additional domain delays (beyond the communication domain) in children already diagnosed with bilateral sensorineural hearing loss (63).

3. The ASQ and ASQ:SE authors should consider improving their tools by making the “overall” section measurable and evaluating how these questions affect validity.

4. Up-to-date ASQ and ASQ:SE norming and validation studies are needed across Scandinavia.

5. Community, statewide and nationwide initiatives help to better implement and sustain broad-band developmental and social-emotional screening (28-30).

References

1. http://pediatrics.aappublications.org/content/127/6/1034.full?sid=235b78be-17e1-40eb-31b-ec5bca3dbd0f


12. OECD Family Database – OECD Social Policy Division – Directorate of Employment,


71. Pontoppidan MW. Translation of the Danish ASQ:SE and Use as an Outcome Measure in Intervention Studies with Infants in Denmark. 2016.


89. Kristensen IH. Video feedback promotes relations between infants and vulnerable first-time mothers. BMC Pregnancy and Childbirth. 2017;17(1). doi:10.1186/s12884-017-1568-1


Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram: ASQ and ASQ:SE use in Denmark, Norway, Sweden and the United States.

Figure 2. How are different countries using or researching the ASQ and ASQ:SE based on article category?
Figure 3. What setting is the ASQ and ASQ:SE being used or researched based on country?

Table 1. Article category and categorization criteria for ASQ and ASQ:SE studies in the United States and Scandinavia (Denmark, Norway and Sweden).

<table>
<thead>
<tr>
<th>Article category</th>
<th>Categorization Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Universal screening on a general or at-risk population article</td>
<td>Article includes analysis of a clearly defined target population, which could be a general population or at-risk group of children with biological and/or environmental conditions that represent developmental-behavioral risk factors and One of the article’s objectives is to systematically screen children for suspected developmental delays and/or social-emotional/behavioral problems and Article reports the entire screening tool’s results with specific cutoffs. Trained professionals or para-professionals (not parents or caregivers) interpreted the screening tool results.</td>
</tr>
<tr>
<td>A. Primary care clinic in a medical setting</td>
<td></td>
</tr>
<tr>
<td>B. Developmental-behavioral, psychiatric or subspecialty clinic/hospital in a medical setting</td>
<td></td>
</tr>
<tr>
<td>C. Nurse home visit setting</td>
<td></td>
</tr>
<tr>
<td>D. Childcare or early learning/preschool setting</td>
<td></td>
</tr>
<tr>
<td>E. Social service setting</td>
<td></td>
</tr>
<tr>
<td>F. Cross-sector setting or birth cohort</td>
<td></td>
</tr>
<tr>
<td>2. Psychometric article</td>
<td>Article includes information on the translation and adaptation process with preliminary reliability data and/or Article includes information on the translation and adaptation process and explicitly discussed or presented data on the feasibility of implementation in a new context or culture and/or Article collects the entire screening tool results in the context of early identification. The sample of participants, reported preliminary prevalence rates or prevalence rates on at least a small sub-sample (≥100) of the population is described.</td>
</tr>
<tr>
<td>A. Primary care clinic in a medical setting</td>
<td></td>
</tr>
<tr>
<td>B. Developmental-behavioral, psychiatric or subspecialty clinic/hospital in a medical setting</td>
<td></td>
</tr>
<tr>
<td>C. Nurse home visit setting</td>
<td></td>
</tr>
<tr>
<td>D. Childcare or early learning/preschool setting</td>
<td></td>
</tr>
<tr>
<td>E. Social service setting</td>
<td></td>
</tr>
<tr>
<td>F. Cross-sector setting or birth cohort</td>
<td></td>
</tr>
</tbody>
</table>
3. Intervention/exposure research article

A. Primary care clinic in a medical setting

B. Developmental-behavioral, psychiatric or subspecialty clinic/hospital in a medical setting

C. Nurse home visit setting

D. Childcare or early learning/preschool setting

E. Social service setting

F. Cross-sector setting or birth cohort

4. Review or policy-related article

It is especially important that the article uses the entire screening tool results (e.g., all 5 domains on the ASQ or all the scored items on the ASQ:SE) to examine or monitor the differences in child development between an intervention (or exposure) group and a control group.

Article scientifically reviews, or provides a professional organization or country-specific policy statement about developmental-behavioral surveillance or screening (and reviews empirical data about the ASQ and/or ASQ:SE).

Other inclusion or exclusion criteria for all categories: 1) For ASQ articles, all 5 developmental domains (communication, gross motor, fine motor, problem-solving and personal-social) needed to be reported to maintain consistency with the ASQ user guide’s reported psychometric and feasibility properties. 2) Articles using other abbreviated versions of the ASQ (e.g. using less than 6 items per domain) and/or ASQ:SE (e.g., haphazardly omitting items that were meant to be scored) were excluded because abbreviated versions are assumed to be psychometrically unsound. 3) The correct age-interval ASQ and/or ASQ:SE should be used. The ASQ-3 and ASQ:SE-2 Users’ Guides and the the American Academy of Pediatrics’ Committee on the Fetus and Newborn (AAP COFN) recommend the age-adjusted interval be administered for preemies born under 37 weeks gestational age. 4) The population under study needed to be exclusively focused on the target countries (Denmark, Sweden, Norway and/or USA) being studied. 5) Articles needed to report or discuss empirical data related to the ASQ and/or ASQ:SE.
Records identified through database searching (n = 1,689)

Additional records identified through other sources (n = 7)

Records after duplicates removed (n = 1,425)

Records screened (n = 1,425)

Records excluded (n = 1,207)

Full-text articles assessed for eligibility (n = 218)

Studies included in qualitative synthesis (n = 127)

Full-text articles excluded, with reasons (n = 91)
39 Study not done in target countries
14 ASQ and/or ASQ:SE used but no empirical data
14 Duplicates
11 Less than 5 ASQ domains used
6 Mixed countries including target country
4 Used abbreviated ASQ:SE and/or ASQ domains
2 No use or discussion of ASQ and/or ASQ:SE
1 ASQ and/or ASQ:SE used but data analysis used combined/unclear results
Figure 2. How are different countries clinically using or researching the ASQ and ASQ:SE based on article category?

153x86mm (96 x 96 DPI)
Figure 3. What setting is the ASQ and ASQ:SE being clinically used or researched based on country?

162x99mm (96 x 96 DPI)
<table>
<thead>
<tr>
<th>Study identification &amp; setting</th>
<th>Overall level of evidence</th>
<th>Comments about study design, objectives, methods (including population)</th>
<th>Study appraisal</th>
<th>Comments about interventions and outcomes (results and conclusions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen, 2010. Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Descriptive observational design. Objective: to describe the Enhancing Developmentally Oriented Primary Care (EDOPC) project’s impact on developmental and social-emotional screening. Methods: Project involved 336 (1-hour, onsite) trainings at 164 primary care clinics throughout Illinois from 2005 to 2006. There were 2873 unique participants and 165 trainings about developmental (ASQ) and social-emotional (ASQ:SE) screening and referring as needed to EI and other community resources for children 0 to 3 years. Selection bias may have been an issue with the primary care practices that chose to participate in the EDOPC project. It is possible that these practices were more enthusiastic about implementing screening tools.</td>
<td>(+) Randomized design with control group being parents who completed the ASQ with non-standardized conditions. Objectives: to examine whether the reproducibility of the ASQ scores differed when it was administered in the waiting room of a clinic compared with “standardized conditions”. Methods: Hospital-based, resident-staffed clinic, urban population in New Haven, Connecticut. 131 English or Spanish-speaking parents of 18-36-month-old children who completed the ASQ in the waiting rooms, then were randomized to repeat the ASQ in waiting room or under standardized conditions which included (1) a private, quiet, distraction-free room, (2) ASQ-specific toys, (3) a trained facilitator who was knowledgeable about the ASQ and could assist with possible literacy barriers.</td>
<td>Intervention = ASQ and ASQ:SE trainings. EDOPC project enhanced confidence and intent to screen among a large group of Illinois primary health care providers. Among a sample of primary care sites at which chart reviews were conducted, the EDOPC project increased developmental (ASQ) screening rates to the target of 85% of patients at most sites and increased social-emotional (ASQ:SE) screening rates to the same target rate in nearly half of the participating practices. Study clearly showed that ASQ and ASQ:SE screening rates increased at a statewide level due to the EPDOC project; however, EU/ECSE linkage and eligibility rates were not measured before and after the EPDOC project was implemented. Nevertheless, this study answered its objective.</td>
</tr>
<tr>
<td>Antonio, 2014. Medical setting, primary care clinic (general population)</td>
<td>(+)</td>
<td>Randomized design with control group being parents who completed the ASQ with non-standardized conditions. Objectives: to examine whether the reproducibility of the ASQ scores differed when it was administered in the waiting room of a clinic compared with “standardized conditions”. Methods: Hospital-based, resident-staffed clinic, urban population in New Haven, Connecticut. 131 English or Spanish-speaking parents of 18-36-month-old children who completed the ASQ in the waiting rooms, then were randomized to repeat the ASQ in waiting room or under standardized conditions which included (1) a private, quiet, distraction-free room, (2) ASQ-specific toys, (3) a trained facilitator who was knowledgeable about the ASQ and could assist with possible literacy barriers.</td>
<td>(+) Randomized design with control group being parents who completed the ASQ with non-standardized conditions. Objectives: to examine whether the reproducibility of the ASQ scores differed when it was administered in the waiting room of a clinic compared with “standardized conditions”. Methods: Hospital-based, resident-staffed clinic, urban population in New Haven, Connecticut. 131 English or Spanish-speaking parents of 18-36-month-old children who completed the ASQ in the waiting rooms, then were randomized to repeat the ASQ in waiting room or under standardized conditions which included (1) a private, quiet, distraction-free room, (2) ASQ-specific toys, (3) a trained facilitator who was knowledgeable about the ASQ and could assist with possible literacy barriers.</td>
<td>Intervention = ASQ screenings. Parents completed the ASQ in the waiting room without help or under “standardized conditions”. No statistically significant differences between intra-class correlation coefficients in waiting room versus standardized conditions in any ASQ domain. In a high-risk, economically disadvantaged population, 25.9% of children failed the ASQ (which could be perceived as a benefit but the potential harms (and costs) of screening were not investigated. The ASQ, when completed by parents in the waiting room without assistance, is reliable compared with standardized conditions, indicating the ASQ can be used to screen children for developmental delay in the waiting room of pediatric practices.</td>
</tr>
<tr>
<td>Arunyanart, 2012. Medical Setting, primary care (general population)</td>
<td>(-)</td>
<td>Cross sectional survey design. Objectives: to determine compliance with AAP screening recommendations: (1) developmental screening at 9, 18, and 24 or 30 months; (2) screening when concerns are raised at a surveillance visit; and (3) autism screening at 18 and 24 months and to examine pediatrician and practice characteristics associated with compliance Methods: 406 pediatricians (all AAP fellows) from 6 US states completed a 38-item web-based questionnaire. Selection bias was a strong concern as only 10% (408 out of 4,100) of pediatricians responded to this survey and only 281 (6.9% of) pediatricians had completed all 38 questions.</td>
<td>(-) Cross sectional survey design. Objectives: to determine compliance with AAP screening recommendations: (1) developmental screening at 9, 18, and 24 or 30 months; (2) screening when concerns are raised at a surveillance visit; and (3) autism screening at 18 and 24 months and to examine pediatrician and practice characteristics associated with compliance Methods: 406 pediatricians (all AAP fellows) from 6 US states completed a 38-item web-based questionnaire. Selection bias was a strong concern as only 10% (408 out of 4,100) of pediatricians responded to this survey and only 281 (6.9% of) pediatricians had completed all 38 questions.</td>
<td>Intervention = ASQ screening and 2008 Family Survey. Parental stress levels and “low feelings of self as a caregiver” were significantly related to more social-emotional problems (concerning ASQ:SE scores) when compared to children of parents with high feelings of self as a caregiver. Findings “suggest that daily living routine, parent stress levels, and parental self-efficacy play an influential role in the development of social-emotional behavior problems in very young children.” Author gave 8 recommendations for parents, communities, state and federal policy makers, and future studies.</td>
</tr>
<tr>
<td>Bergman, 2009. Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Descriptive observational design. Objective: to evaluate the feasibility and acceptance of a new model for well-child care in a large health maintenance organization. Methods: involved about 100,000 children and families at Kaiser Permanente (a HMO) in the metropolitan area of Denver, Colorado. An Internet-based tool, the Child Health and Development Interactive System (CHADIS) used the online ASQ-2 (among many other screening tools) 2 weeks prior to well-visits for children ages 0 to 3 years. Study focus was about implementation, not accuracy of online prevvisit ASQ results.</td>
<td>(+) Descriptive observational design. Objective: to evaluate the feasibility and acceptance of a new model for well-child care in a large health maintenance organization. Methods: involved about 100,000 children and families at Kaiser Permanente (a HMO) in the metropolitan area of Denver, Colorado. An Internet-based tool, the Child Health and Development Interactive System (CHADIS) used the online ASQ-2 (among many other screening tools) 2 weeks prior to well-visits for children ages 0 to 3 years. Study focus was about implementation, not accuracy of online prevvisit ASQ results.</td>
<td>Intervention = CHADIS (with pre-visit online ASQ screening). 75% of parents thought that online pre-visit screening &amp; assessment process “improved” or “very much improved” their well-visit experience. However, 12% of parents found online pre-visit screening &amp; assessment process “somewhat or very difficult to use” in this lower risk population whose parents are employed and receive healthcare at a HMO. Authors demonstrated the feasibility of a new well-child care model that used parent-report pre-visit screening and alternative visit types to tailor care to the needs of the family.</td>
</tr>
<tr>
<td>Reference</td>
<td>Setting</td>
<td>Design</td>
<td>Objective</td>
<td>Interventions</td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
</tr>
<tr>
<td>Berry, 2014</td>
<td>Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Descriptive observational design. Objective: to increase the financing and delivery of preventive developmental services for children birth to age 3 years in the state of Illinois. Methods: Enhancing Developmentally Oriented Primary Care (EDOPC) project involved 336 1-hour, onsite trainings at 164 primary care practices throughout Illinois from 2005 to 2006. There were 2873 unique participants and 165 trainings about developmental (ASQ) and social-emotional (ASQ:SE) screening and referring as needed to EI and other community programs for children 0 to 3 years. Selection bias concern: were the primary care practices that chose to participate in the EDOPC project more enthusiastic about implementing the ASQ and ASQ:SE?</td>
<td>Intervention = ASQ and ASQ:SE screenings. EDOPC project staff determined best practices for developmental (ASQ), social-emotional (ASQ:SE) and other screening, and for facilitating referrals for children with suspected delays and/or psychosocial risk factors. Implementation used a team-based approach with a leader. Each staff member was assigned to do specific tasks. Plan-Do-Study-Action cycles were used to improve work flow, along with screening rates over time. EI/ECSE linkage and eligibility rates were not measured before and after the EPDOC project was implemented. Nevertheless, this study did answer its clearly stated objective.</td>
</tr>
<tr>
<td>Briggs, 2012</td>
<td>Medical setting, primary care clinic (general population)</td>
<td>(+)</td>
<td>Prospective cohort design with intervention and control groups. Objectives: to describe a program designed to identify the social-emotional status of young children in the pediatric clinic with the ASQ:SE and to assess the effect of interventions by a colocated psychologist on changes in ASQ:SE scores over time. Methods: Study analyzed ASQ:SE scores completed on children ages 6 to 36 months, to determine if they were at risk for social-emotional problems. Probability of remaining at risk over time was compared between subjects receiving intervention by the psychologist, and those who declined intervention. Logit specifications were used in multivariate comparisons to control for a set of covariates.</td>
<td>Intervention = ASQ:SE screenings. 3169 children screened; 711 (22.4%) scored ≥ the refer/at-risk cutoff. Among 711 at-risk children, 170 were rescreened when older. Children who received intervention from the psychologist showed significant improvement on ASQ:SE scores compared with those whose parents declined intervention (P = .01). Implication of the results indicate that the colocation of a psychologist in primary care clinic promotes the ability to effectively address young children’s social-emotional development. However, EI/ECSE linkage and eligibility rates were not measured. There was no reference standard or diagnostic evaluation</td>
</tr>
<tr>
<td>Briggs, 2014</td>
<td>Medical setting, primary care clinic (general population)</td>
<td>(+)</td>
<td>Quasi-experimental, longitudinal design. Objective: to describe the Healthy Steps (HS) program and the moderating effect of this program on the relationship between reported caregiver childhood trauma &amp; child social-emotional development. Methods: 124 children assessed at 36 months. Children of mothers with childhood trauma had higher/worse ASQ:SE mean scores than children of mothers without childhood trauma (75.9 vs 35.9; P&lt;.0001). Differences in adjusted mean ASQ:SE scores between children of mothers with/without childhood trauma were more apparent in the comparison group (90.4 vs 28.3) than HS (44.5 vs 28.2; P&lt;.001).</td>
<td>Intervention = ASQ:SE screening. Of the 132 patients identified in the Cardiothoracic Surgery database and at discharge from the hospital, a total number of 106 infants were reviewed. A genetic syndrome was identified in 23.4% of the population. Neuroimaging abnormalities were identified in 21.7% of the cohort with 12.8% having visibly severe insults. 23 (26.7%) at-risk children received first-time referrals for EI services, 16 (13.8%) received referrals for new services in addition to their existing ones. Utilization of the ASQ (and other existing resources) in collaboration with established programs can ensure targeted neurodevelopmental follow-up for children with complex congenital heart disease.</td>
</tr>
<tr>
<td>Brown, 2012</td>
<td>Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Cross-sectional observational design. Objective: to estimate the prevalence of positive screens for social-emotional problems with the ASQ:SE among preschool-aged children in a low-income clinical population and to explore the family context and receptivity to referrals to help guide development of interventions. Methods: 254 parents of 3- and 4-year-old children at 2 urban primary care clinics that served low-income children in Cincinnati, Ohio. Study done between June and November 2010 which is not the busy season in pediatric clinics. No reference standard or diagnostic testing done on children with positive/concerning ASQ:SE screening results.</td>
<td>Intervention = ASQ:SE screening. 24% (95% CI, 16.5%-31.5%) of children screened positive on the ASQ:SE. Among those, 45% had a parent with depressive symptoms, and 27% had no nonparental child care. Among parents of children who screened positive on the ASQ:SE, 79% reported they would welcome or not mind a referral to a counselor or psychologist; only 16% reported a prior referral. Opportunities exist to improve the prevention of social-emotional problems and improve rates of children receiving ECSE services.</td>
</tr>
<tr>
<td>Chiu &amp; DiMarco, 2010</td>
<td>Medical setting, primary care clinic (at risk population)</td>
<td>(+/-)</td>
<td>Prospective, comparative design pilot study. 20 homeless mothers and their 21 children who were between 4 to 60 months. Inclusion criteria for mothers: English speaking, aged 18 years or older, residing in homeless shelter, not under the influence of drugs or alcohol during shelter stay, not presently victims of domestic violence and not directly released from mental facility. No diagnostic testing done on children with concerning ASQ screening results. Denver II screening test is out-of-date, psychometrically unsound and is not a reference/cold standard test.</td>
<td>Interventions = ASQ and Denver II screening. Pediatric nurse practitioners administered the Denver Developmental Screening Test II and parents completed the ASQ. The percentage agreement between 2 tools was strongest in gross motor (95%) and personal-social domains (95%) but weakest in the communication/language (67%) domains. ASQ screening was found to be helpful and acceptable by homeless mothers. Study is out-of-date and its clinical impact in universal screening settings is unclear.</td>
</tr>
<tr>
<td>Chorna, 2016</td>
<td>Medical setting, neurodevelopmental and pediatric cardiology clinic (at risk population)</td>
<td>(+)</td>
<td>Retrospective observational design. Objective: to provide a feasible and responsible utilization of the existing infrastructure and personnel, and to develop and implement a developmental-behavioral screening and assessment program dedicated to children with congenital heart disease. Methods: Trained testers administered ASQ-3 over the phone to parents of all referred children at least once between 6 and 12 months’ corrected age. At 18 months’ corrected age, all children were scheduled in the NICU Follow-Up Clinic for a visit with standardized neurological exams, the Bayley III, and multidisciplinary therapy evaluations. Of 132 patients identified in the Cardiotoracic Surgery database and at discharge from the hospital, 106 infants were reviewed.</td>
<td>Interventions = ASQ screening and neuro-developmental diagnostic testing. Of the 132 patients identified in the Cardiotoracic Surgery database and at discharge from the hospital, a total number of 106 infants were reviewed. A genetic syndrome was identified in 23.4% of the population. Neuroimaging abnormalities were identified in 21.7% of the cohort with 12.8% having visibly severe insults. 23 (26.7%) at-risk children received first-time referrals for EI services, 16 (13.8%) received referrals for new services in addition to their existing ones. Utilization of the ASQ (and other existing resources) in collaboration with established programs can ensure targeted neurodevelopmental follow-up for children with complex congenital heart disease.</td>
</tr>
</tbody>
</table>
Duffner, 2012. Medical setting, metabolic sub-specialty clinic (at-risk population)

Longitudinal prospective cohort study. Objectives: To assess the utility of a telephone-based interview system in providing monitoring of the developmental and functional status of children with both positive newborn screens for Krabbe disease and low galactocerebrosidase activity on confirmatory testing, and to determine whether this approach provides improved compliance with follow-up compared with formal neuropsychological testing. Methods: 17 infants (8 males, 9 females) with low galactocerebrosidase activity (as detected by New York State newborn screening program). Consenting families sent age-appropriate ASQ-2 to complete before telephone interviews at ages 4, 8, 12, 18 and 24 months. ASQ screenings and diagnostic instruments used at 12 and 24 months. Monitoring until <24 months means follow up might not have been complete enough to detect late-onset Krabbe disease.

Earls, 2009. Medical setting, primary care clinics (general population)

Prospective observational cohort design. Objectives: to determine the percentage of children screened with the ASQ and whether screening rates improved with time, to observe patterns and trajectories for children identified at risk in 1 or more of the 5 ASQ domains, and to examine referral rates and physician referral patterns. Methods: 526 children followed from August 2001 through November 2003 from primary care practices that were participating in the North Carolina All Children's Health and Development (ABCD) project. 1143 parent-completed ASQs were administered, scored and interpreted at the 6, 12, 18 or 24, 36, 48 and 60 month well-child visits. Follow up EI referral and statewide EI eligibility rates were measured.

Eom, 2014. Medical setting, pediatric neurology sub-specialty clinic (at-risk population)

Prospective observational design. Objectives: to assess the value of the ASQ-3 and its validity for determining previously unidentified (‘actionable’) problems in children with epilepsy. New and existing patients with epilepsy recruited from a hospital-based epilepsy center. The parent completed screening for developmental delays (ASQ-3, 0–66 mo), autism (Modified Checklist for Autism in Toddlers [M-CHAT], 16–30 mo), communication social communication [Social Communication Questionnaire [SCQ], 24y), and social-emotional skills (Strengths and Difficulties Questionnaire [SDQ], 4–17y). 236 children screened overall (58% males, 42% females; mean age [SD] 6y 7mo [4y 6mo]). Need for anti-seizure medicine(s) could have been an unaccounted confounding factor.

Fisher, 2012. Medical setting, pediatric neurology sub-specialty clinic (at-risk population)

Prospective observational design. Objective: Study assessed yield of routine screening for developmental delays and autism in a hospital-based, tertiary care center. Methods: Parents completed ASQ and Modified Checklist for Autism in Toddlers (M-CHAT) for 65 children (average age = 2.5 y; 58% boys) who came from a socio-economically and racially diverse population in the Chicago metropolitan area. Only 1 family refused to participate in the screening study. No reference/gold standard or diagnostic testing was done but appropriate referrals for further evaluations or services were made for 16 children with new concerns identified by the ASQ or other screening tools. Need for anti-seizure medicine(s) could have been an unaccounted confounding factor.

Goldberg, 2014. Medical setting, neurodevelopmental and cardiology clinic (at risk population)

Prospective observational design. Objectives: To measure neurodevelopment at 3 years of age in children with single right-ventricle anomalies and to assess its relationship to Norwood shunt type, neurodevelopment at 14 months of age, and patient and medical factors. Methods: All subjects in the Single Ventricle Reconstruction Trial who were alive without cardiac transplant were eligible for inclusion. 203 children screened with ASQ and other measures of behavior and quality of life at 3 years. Medical history, including measures of growth, feeding, and complications, was assessed through annual review of the medical records and phone interviews. Bayley Scales of Infant Development, 2nd Ed. (BSID-II) scores at 14 months evaluated as predictors.

Grossman, 2010. Medical setting, pediatric emergency room (at-risk population)

Cross-sectional study design. Objectives: To determine whether screening children in an urban pediatric emergency department would lead to identification of previously undiagnosed developmental delay. Methods: Families in an urban public hospital pediatric emergency department with children 6 to 36 months and no history of a developmental delay or receiving EI services. 138 children screened with ASQ. Mean age 18.9 months; 51.5% female; 56.8% of mothers high-school graduates; 59.9% immigrants; 75.4% Latino. 21% no primary care provider.

Intervention = ASQ screenings. 17 patients were enrolled; 16 were assessed at age 12 and 18 months, and 15 were assessed at age 24 months. Scores were within the normal range on ASQ-2 and tests of functional status, with the exception of expressive language. Only 7 patients completed the Bayley Scales of Infant and Toddler Development, Third Edition assessments; all their scores were in the normal range.

Prospective observational design. Objectives: to determine the percentage of children screened with the ASQ and whether screening rates improved with time, to observe patterns and trajectories for children identified at risk in 1 or more of the 5 ASQ domains, and to examine referral rates and physician referral patterns. Methods: 526 children followed from August 2001 through November 2003 from primary care practices that were participating in the North Carolina All Children's Health and Development (ABCD) project. 1143 parent-completed ASQs were administered, scored and interpreted at the 6, 12, 18 or 24, 36, 48 and 60 month well-child visits. Follow up EI referral and statewide EI eligibility rates were measured.

Intervention = ASQ screenings. Number of screenings per child steadily increased over the first 2 to 2.5 years of the project. Providers/practices improved office-based ASQ implementation using a team-based approach. Children were referred to EI at earlier ages. According to surveyed providers, there is a greater likelihood of and EI referral for gross-motor concerns, followed by communication concerns. 100% of providers would refer if multiple developmental domains indicated a suspected delay. Providers were less certain about needing EI referrals for fine-motor, problem-solving, and social-emotional skills. Follow up research, statewide ASQ screening rates increased fivefold, EI referral rates increased fourfold, and the % of children who qualified for EI services in North Carolina increased from 3.0% to 4.3%.

Intervention = ASQ and other screening tools. 176 out of 236 children (75%) had established epilepsy diagnoses and 60 (25%) were patients with new-onset epilepsy. Positive findings by test were 82% (ASQ-3), 54% (M-CHAT), 15% (SCQ), and 58% (SDQ). Findings were actionable in 46 children (20%): 18% of findings in children with established epilepsy and 23% in patients with new-onset epilepsy. Of the 46 children for whom further referrals were made, the parents of 28 (61%) pursued further evaluations. Findings supported systematic ASQ screening of comorbidities (e.g, suspected developmental delays, early signs of autism spectrum disorders) for children with epilepsy.

Intervention = ASQ, and MCHAT screenings. 49 (75%) were established epilepsy patients, and 16 (25%) were new patients. Developmental findings (already receiving services) were confirmed in 32/55 (58%) and actionable steps were taken in 17 (31%) children who required further evaluation. The yield of routine ASQ and autism screening of children in a tertiary center was sufficiently high to consider screening of all children with epilepsy.

Intervention = ASQ screening and many other screening and diagnostic instruments. Children with single right-ventricle anomalies have impaired neurodevelopment at 3 years of age. Lower ASQ scores are associated with medical morbidity, and lower BSID-II scores but not with shunt type. However, because only a modest percentage of variation in 3-year neurodevelopmental outcome could be predicted from early measures, all children with single right-ventricle anomalies should be followed longitudinally to improve recognition of delays.

Intervention = ASQ screenings. Almost 30% of 6- to 36-month-old children from an urban pediatric emergency department (without prior developmental concerns) screened positive for suspected delays, suggesting the utility of routine broad-band developmental screening in this setting. Pediatric emergency department use alone may be an indication for ASQ screening. Further study is needed for feasibility of routinely screening for delay in emergency rooms. There was no reference/gold standard or diagnostic testing.
<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guevara, 2013</td>
<td>Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Randomized controlled, parallel-group trial. Objectives: To determine the effectiveness of developmental screening on the identification of developmental delays, EI referrals and EI eligibility. Methods: 2-year study in 4 urban pediatric practices. Children eligible if they were &lt;30 months old, term, without congenital malformations or genetic syndromes, not in foster care, and not EI-enrolled. Children randomized to 1 of 3 groups: (1) ASQ-2 and Modified Checklist for Autism in Toddlers (M-CHAT) screening with office staff assistance, (2) ASQ-2 and M-CHAT screening without office staff assistance, or (3) developmental surveillance only (using age-appropriate milestone checklists) at well-visits. Outcomes were assessed using an intention-to-treat analysis. 2103 children enrolled (most African American with family incomes &lt; $30,000/year). Parents and providers were not blinded to children’s participation in 1 of the 3 groups.</td>
<td>Interventions = ASQ and autism screening. In the first 3 years of the Quick Peek Program, 1,150 children were screened with the ASQ and MCHAT-R and 589 children were lost to follow-up at 1 month. Of at-risk children contacted by 1-month follow-up, 380 (88%) had arranged recommended evaluations and services and 50 (12%) did not follow through. Models need to include training a variety of community providers (not just pediatricians) including child care workers to conduct developmental screening face-to-face with the children and their caregivers. Reference/gold standard or diagnostic testing was not done.</td>
<td>Interventions = ASQ and autism screening. Children in screening arms 1 and 2 were more likely to be identified with delays (23.0% and 26.8% vs 13.0%; P&lt;.001), referred to EI (19.9% and 17.5% vs 10.2%; P&lt;.001), and eligible for EI services (7.0% and 5.3% vs 3.0%; P&lt;.001) than children in the surveillance arm. Children in the screening arms incurred a shorter time to identification, early intervention (EI) referral, and EI evaluation than children in the surveillance arm. Children who participated in a developmental screening program were more likely to be identified with developmental delays, referred to EI, and eligible for EI services in a timelier fashion than children who received surveillance alone. Study supported the 2006 AAP policy statement. Future efforts should focus on care coordination between EI and medical homes. Potential harms and costs were not investigated or measured.</td>
</tr>
<tr>
<td>Hardy, 2015</td>
<td>Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Prospective observational design. Objective: to examine the ASQ-3’s ability to identify children at-risk for autism. Authors looked at ASQ-3 scores of children who screened positive on the Modified Checklist for Autism in Toddlers-Revised (M-CHAT-R), children who continued to screen positive on the MCHAT-R Follow-up Interview, and children diagnosed with autism spectrum disorders. Methods: 2848 toddlers, aged 16 to 30 months, screened with ASQ-3 and Modified Checklist for Autism in Toddlers-Revised (MCHAT-R) across 20 pediatric sites and 20 pediatricians across Connecticut. Data collected over 4-year period (2009-2013). Children who screened positive on the MCHAT-R and its Follow-up Interview were offered a diagnostic evaluation. A group of 207 Spanish speakers were included, using the Spanish MCHAT-R and ASQ-3.</td>
<td>Interventions = ASQ-R and autism screening. In the first 3 years of the Quick Peek Program, 1,150 children were screened with the ASQ and MCHAT-R and 589 children (51%) were found to be “at risk.” For at-risk children, 159 (21%) of at-risk children were lost to follow-up at 1 month. Of at-risk children contacted by 1-month follow-up, 380 (88%) had arranged recommended evaluations and services and 50 (12%) did not follow through. Models need to include training a variety of community providers (not just pediatricians) including child care workers to conduct developmental screening face-to-face with the children and their caregivers. Reference/gold standard or diagnostic testing was not done.</td>
<td>Interventions = ASQ and MCHAT-R screenings. Using the monitor and/or refer cutoffs on any domain, ASQ-3 identified 87% of children who screened positive on M-CHAT-R with Follow-up Interview and 95% (20/21) of those diagnosed with autism. Monitor and/or refer on the communication domain alone identified 95% of children diagnosed with autism. Scores below monitor cutoff on communication domain indicates initial concern requiring autism-specific follow-up. All screened children did not get reference/gold standard testing. Nevertheless, if results are confirmed with a large enough sample to separately examine toddlers of different ages and cultural backgrounds, a 2-stage screening strategy may be feasible, with autism screening reserved for those with a positive ASQ-3. Possibly very impactful results and conclusions but further study is needed.</td>
</tr>
<tr>
<td>Harris &amp; Norton, 2016</td>
<td>Medical setting, primary care clinics (general population)</td>
<td>(+)</td>
<td>Prospective observational design. Objective: to create and implement a “Quick Peek” model to provide free community-based developmental screening in order to improve early access to care among a primarily underserved population. Methods: Free, bilingual clinics provided within underserved communities, targeting 1150 children screened between 1 and 5 years old. The ASQ-3 was conducted interactively with the child, parent/guardian, and screener, as well as the Modified Checklist for Autism in Toddlers-Revised (MCHAT-R) when applicable for child’s age (from 15 to 30 months). Clinics based in low-income cities in New Jersey with a large percentage of racial minority, Hispanic/Latino, and/or Spanish-speaking households.</td>
<td>Intervention = ASQ-3 screening and MCH-CHAT-R. Children selected for this intervention were reached by phone and asked to attend a scheduled appointment in the clinic in order to complete the screening process. The purpose of this intervention was to increase awareness of the importance of early identification of developmental delays among children in low-income communities. Children who screened positive on the ASQ-3 and/or MCH-CHAT-R were referred to the child’s pediatrician or primary care provider for follow-up evaluation and assessment. Reference/gold standard or diagnostic testing was not measured.</td>
<td>Intervention = ASQ and MCHAT-R screenings. Using the monitor and/or refer cutoffs on any domain, ASQ-3 identified 87% of children who screened positive on M-CHAT-R with Follow-up Interview and 95% (20/21) of those diagnosed with autism. Monitor and/or refer on the communication domain alone identified 95% of children diagnosed with autism. Scores below monitor cutoff on communication domain indicates initial concern requiring autism-specific follow-up. All screened children did not get reference/gold standard testing. Nevertheless, if results are confirmed with a large enough sample to separately examine toddlers of different ages and cultural backgrounds, a 2-stage screening strategy may be feasible, with autism screening reserved for those with a positive ASQ-3. Possibly very impactful results and conclusions but further study is needed.</td>
</tr>
<tr>
<td>Haskett, 2016</td>
<td>Social service setting, homeless shelter community project (at risk population)</td>
<td>(+)</td>
<td>Prospective observational cohort design. Objective: to gain understanding of homeless children’s social–emotional adjustment and their functioning in language, motor, and cognitive skills. Methods: 328 children residing with their parents in one of 11 emergency shelter or transitional housing programs for families who were experiencing homelessness in a central North Carolina county. Child case managers administered the Brigance Early Childhood Screen II and the ASQ-SE in the homeless shelter setting. Data collected by Community Action Targeting Children who are Homeless (CATCH) project. Population was mostly African American.</td>
<td>Intervention = ASQ SE screening and the Brigance Early Childhood Screen II. Developmental scores for overall functioning of the sample were significantly below the norming group, with particularly low functioning in language and communication skills. Parents of 24.8% of the children had substantial concerns about their children’s mental health. Although there are individual differences in adjustment of children experiencing homelessness, results support wide-scale screening and access to early intervention for these vulnerable children. EI/ECSE eligibility rates or reference/gold standard or diagnostic testing was not measured.</td>
<td>Intervention = ASQ-3 screening and the Brigance Early Childhood Screen II. Developmental scores for overall functioning of the sample were significantly below the norming group, with particularly low functioning in language and communication skills. Parents of 24.8% of the children had substantial concerns about their children’s mental health. Although there are individual differences in adjustment of children experiencing homelessness, results support wide-scale screening and access to early intervention for these vulnerable children. EI/ECSE eligibility rates or reference/gold standard or diagnostic testing was not measured.</td>
</tr>
<tr>
<td>Hernandez-Mekonnen, 2016</td>
<td>Medical setting for immigrants (at risk population)</td>
<td>(+)</td>
<td>Prospective observational cohort design. Objective: to examine the rates of positive screening for developmental delay and EI service utilization in an unauthorized Mexican immigrant community and factors associated with each. Methods: Interviews conducted in Spanish with Mexican born women receiving maternal health care. Children 12–60 months of age were screened with ASQ. Participants recruited from a Women’s Clinic providing perinatal care to uninsured Spanish-speaking women in Philadelphia, PA. Data collected from June, 2012 to February, 2013. Women born in Mexico, over 18 years old and either pregnant or had a child age 0–66 months were eligible to participate.</td>
<td>Intervention = ASQ screenings. For children with concerning ASQ results, 38% (n = 3) received early intervention (EI) in a timely manner. An additional 26% (n = 17) qualified for further monitoring, and of those, 59% (n = 10) received EI. Women with low health literacy had more than four times the odds of having a child with risk of developmental delay (aOR 4.4; 95% CI 1.3–15.4). Developmental delay was associated with low maternal health literacy in unauthorized Mexican immigrants. Despite higher EI use rates, care coordination is critically important for these at-risk children.</td>
<td>Intervention = ASQ screenings. For children with concerning ASQ results, 38% (n = 3) received early intervention (EI) in a timely manner. An additional 26% (n = 17) qualified for further monitoring, and of those, 59% (n = 10) received EI. Women with low health literacy had more than four times the odds of having a child with risk of developmental delay (aOR 4.4; 95% CI 1.3–15.4). Developmental delay was associated with low maternal health literacy in unauthorized Mexican immigrants. Despite higher EI use rates, care coordination is critically important for these at-risk children.</td>
</tr>
<tr>
<td>Authors</td>
<td>Design</td>
<td>Objectives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hix-Small &amp; Marks, 2007. Medical setting, primary care clinics (general population)</td>
<td>(+) Prospective pre-post cohort design. Objectives: to investigate the effectiveness and costs of incorporating a parent-completed developmental screening tool, the Ages and Stages Questionnaire, into the 12- and 24-month well-child visits under &quot;real-world&quot; conditions, using a combined in-office and mail-back data collection protocol. Methods: 1428 caregivers and children presenting for the 12- or 24-month well-visit between April 2005 and March 2006 participated. Children already identified with EI-eligible delays or disorders were excluded. 18 board-certified pediatricians and 2 pediatric nurse practitioners (from 3 different clinics) acted as secondary participants. Pediatricians blinded to ASQ results when completing their “Pediatrician Developmental Impression” immediately after the well-visit (i.e., was the child typical vs borderline vs delayed? If borderline or delayed, then in which of 5 domains?). Patients with suspected delays on the ASQ or Pediatrician Developmental Impression results were all referred to an EI agency. Participants were from Oregon (38% low income, 72% white, 14% Hispanic, 9.3% Spanish speaking, 7% multiracial, 5% Asian/Pacific Islander, 1% American Indian, 2% African American).</td>
<td>Intervention = ASQ screening routinely at 12 and 24 months. ASQ found to be low cost and feasible for routine use at well-child visits. The 54% ASQ return rate, though acceptable under study conditions, called for alternative implementation strategies (like online pre-visit screening) rather than a post-visit, mail-back protocol. Because of the ASQ, EI referral rates increased 224% (5.7-fold increase at 12-months and 2.9-fold at 24 months) compared to pediatricians’ developmental impression (surveillance alone). Pediatrician referrals (blind to ASQ results) were highly likely to qualify for EI services (96%); however, they accounted for only 42% of total EI referrals, highlighting the need for routine ASQ screening. Relying on the pediatrician’s impression alone (which relied on yes/no items extrapolated from the Denver II), approximately half of the children who qualified for EI services would have been missed. Pediatric-only EI referrals (blind to ASQ results) were significantly predicted by a suspected communication delay and/or gross motor delay. After ASQ implementation at 12 and 24 months, EI referral and EI eligibility services dramatically increased compared to control year.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hunter &amp; Lynch, 2014. Medical setting, primary care clinics (general population)</td>
<td>(+) Prospective cohort design. Objective: to describe the implementation of the Ages and Stages Questionnaire: Social-Emotional (ASQ:SE) in primary care practice by mail when children are 30 months old. Methods: 4-month study where the parents of 30-month-old children were mailed the ASQ:SE. In children who did not pass screening or received a call from a registered nurse for parental-report concerns, short-term clinical outcomes were obtained from the electronic medical record. During the last month of the study, the demographics variables of race and insurance type were analyzed for an association with questionnaire completion by mail. Study was conducted at an academic center-based primary care practice (consisting of 4 locations in Rochester and Kasson, Minnesota) with 870 families from October, 2010 to February, 2011 when pediatric clinics tend to be busier.</td>
<td>Intervention = ASQ:SE screenings. Of the 870 families mailed 30-month ASQ:SE screens, 507 (58.3%) were returned by mail. Out of the children with returned screens, 38 (7.5%) of parents were contacted for either elevated scores or concerning comments and 6 (1.2%) were referred to EI program. Parents of children with government insurance returned the ASQ:SE questionnaire 34.2% (13/38) of the time compared with 65.5% (76/116) of those with private insurance ($P &lt; .001). ASQ:SE screening can be effectively managed in primary care practice by a registered nurse using a follow-up protocol. Mailing the ASQ:SE is likely not an effective way to comprehensively screen most primary care populations. Alternative ASQ:SE implementation strategies (online pre-visit screening) are needed. EI/ECSE eligibility rates or reference/gold standard or diagnostic testing was not measured.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jee, 2010. Medical setting, primary care clinic for foster children (at risk population)</td>
<td>(+) Retrospective observational design. Objectives: to examine: (1) whether systematic social-emotional screening improves detection rates of social-emotional problems, compared to clinical judgment; (2) the relative effectiveness of two validated social-emotional screening tools; and (3) the patterns of social-emotional problems among children in foster care. Methods: Chart review done of children in foster care ages 6 months to 5.5 years: 192 children before and 159 after screening implementation, to measure detection rates for social-emotional problems among children. The ASQ-SE and the ASQ were used in multivariable logistic regression analyses to examine associations between children with social-emotional problems. Conducted in a practice that exclusively sees children in foster care in Rochester, New York.</td>
<td>Intervention = ASQ and ASQ:SE screenings. 24% of children identified with a social-emotional problem with screenings, while provider surveillance detected 4%. Agreement between the ASQ:SE and ASQ ranged from 56% to 75%, when data were stratified by age group. Preschool children were more likely to have a social-emotional problem than toddlers and infants (aOR = 3.4, 95% CI = 1.1–10.8). Systematic ASQ:SE screening significantly increased detection rate for social-emotional problems among foster care children, compared to provider surveillance alone and ASQ screening. EI/ECSE eligibility rates or reference/gold standard or diagnostic testing was not measured.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jee, 2010. Medical setting, primary care clinic for foster children (at risk population)</td>
<td>(+) Pre-post observational cohort study design. Study screened all children in foster care for developmental delays by using the ASQ-2. The baseline detection rate was determined by medical chart review for all children aged 4 to 61 months who were new to foster care (NFC) during a 2-year period. After implementation, caregivers of young children who were NFC or already in foster care (IFC) completed the ASQ at well-child visits. The feasibility of systematic screening (the percentage of ASQs completed among the NFC and IFC groups). Detection of delays compared among the baseline NFC group and the screening-NFC group using bivariate and multivariable logistic regression.</td>
<td>Intervention = ASQ screening. Of 261 visits that occurred after initiation of screening, 251 (96%) visits had a completed ASQ form in the medical chart, demonstrating high feasibility. Among children who were NFC, the detection of developmental delay was higher in the screening than baseline period for the entire population (58% vs 29%; $P &lt; .001), for each age group (infants: 37% vs 14%; toddlers: 89% vs 42%; preschool: 82% vs 44%; all $P &lt; .01), and for all developmental domains. On adjusted analyses, the detection of potential developmental delay in toddler and preschool children was higher among the NFC screening group than the NFC baseline group. Systematic screening with the ASQ was found to be feasible and doubled the detection of children with suspected developmental delays. EI/ECSE eligibility rates or reference/gold standard or diagnostic testing was not measured.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### King, 2010.
**Medical setting, primary care clinics (general population)**

Prospective observational design. Objectives: to assess the degree to which a national sample of pediatric practices could implement American Academy of Pediatrics (AAP) recommendations for developmental screening and referrals, and to identify factors that contributed to the successes and shortcomings of these efforts. Methods: American Academy of Pediatrics (AAP) launched the Developmental Surveillance and Screening Policy Implementation Pilot (D-PIP), which was a 9-month pilot project in which 17 diverse practices (from 15 US states) sought to implement the AAP 2006 developmental screening and surveillance policy statement’s recommendations. Pediatric primary care clinics divided up responsibilities among pediatricians, office staff and actively monitored implementation. Despite these efforts, many practices struggled during busy periods (winter) and times of staff turnover. Quantitative data from chart reviews were used to calculate rates of screening and referral. Qualitative data on practices’ implementation efforts were collected through semi-structured telephone interviews and inductively analyzed to generate key themes. No reference standard/gold or diagnostic testing was done.

Intervention = ASQ and the Parents’ Evaluation of Developmental Status (PEDS) screenings. At the D-PIP project’s conclusion, practices reported screening >85% of patients at 9, 18 and 24-30 months. Most practices were unable or unwilling to adhere to 3 specific AAP recommendations: implementing a 30-month visit; administering a screen after surveillance suggested concern, and submitting simultaneous referrals to medical subspecialists and EI programs. Children at practices that used the PEDS “failed” their screens twice as often as children at practices that used the ASQ. Overall, practices reported referring only 61% of children with failed screens. Referral rates for children with a failed PEDS were far lower than those for a failed ASQ (43% vs 72%; P < .001). Many practices seemed to struggle with tracking EI referrals. Many referred children did not follow up with an EI agency. A diverse sample of practices successfully implemented AAP screening as recommendations. Practices were less successful in generating and tracking EI referrals. More attention is needed on referral care coordination.

### Lehr, 2016.
**Medical setting, primary care clinic – teen parent program (at risk population)**

Prospective observational cohort design. Objective: to understand the association between parenting stress, child age and gender, and risk of developmental delay. Methods: Study done at an adolescent health center which serves mostly low-income youth in New York City who are 10 and 24 years of age and primarily of ethnic minority background. The majority of patients (60.5%) in the Teen Parent Program report living in households receiving Medicaid or not having medical insurance and 74% are of low socioeconomic status (having themselves or their family ever used public assistance, food stamps, Medicaid, WIC, free school lunch, or suffered from food scarcity). Eligible participants were caregiving mothers who: (1) attended a well-visits from November 2010 to April 2014; (2) spoke English; and (3) completed the ASQ-3. No reference standard/gold or diagnostic testing was done after ASQ-3 screenings.

Intervention = ASQ screenings. 32% of children showed increased risk of developmental delay using the ASQ-3. Maternal age did not increase the odds of a delay. Multivariate analyses showed that the adjusted odds ratio of delay was higher among boys (vs girls), in personal-social (aOR = 6.2) and overall (aOR = 2.7). Compared to baseline age group (1–9 months), the odds of delay were higher in age 10–18 months for personal-social (aOR = 5.3); in age 19–31 months for communication (aOR = 13.5) and overall (aOR = 6.3); and in age 32–44 months for communication (aOR = 7.0). The odds of a delay were higher in children of mothers with higher parenting stress in 3 of 5 domains (fine motor, problem solving, personal-social) and overall. Children’s ages 3–4 and boys were associated with increased odds of delay in language and personal-social domains. Beyond ASQ screening, surveillance for highly stressed mothers may be important to broaden risk reduction efforts among young mother–child dyads served in primary care settings.

### Lyman, 2007.
**Cross-sector setting (general population)**

Qualitative survey study. Objective: to seek the feedback of early childhood service providers regarding the usefulness of the ASQ:SE in multicultural populations. Methods: Authors collected feedback from a multidisciplinary panel of professionals about using the ASQ:SE for screening multicultural populations. Structured interviews and focus groups done on professionals working at EI, early learning, preschool (Head Start), home visits, medical and mental health programs.

Intervention = ASQ:SE screenings. There are many complexities inherent in conducting social-emotional or psychosocial screening in cross-cultural contexts. When ASQ:SE is used in a culturally competent manner, it enhances the dialogue between cultures and the multiple meanings of risk and resilience. “ Culturally competent manner” was not measured or clearly defined. Open acknowledgement of ethnocentrism & resulting stereotyping helps cultural competence with the process of ASQ:SE screening. Topic was interesting but the practical value of the study was vague for busy medical clinics.

### Lynch, 2015.
**Medical setting, primary care clinics (general population)**

Prospective cohort design. Objective: to evaluate the effectiveness of mailing the ASQ-3 and Modified Checklist for Autism in Toddlers (MCHAT) at 18 months, with a standardized follow-up process for abnormal results. Methods: Parents of 892 children received by mail the 18-month ASQ and MCHAT from December 2008 to September, 2009. Parents were previously instructed they would need to fill out the ASQ-3 and MCHAT at the 15-month well-visit. A registered nurse scored the questionnaires and, if needed, administered follow up screening and/or a referral to an EI agency and/or a pediatric subspecialist. Children with abnormal ASQ-3 results were all offered additional assessment with a general pediatrician and a trained nurse who administered the Capute Scales (Cognitive Adaptive Test/Clinical Linguistic Auditory Milestone Scale or CAT/CLAMS). Medical record reviews determined clinical outcomes of children who required intervention after the initial screening through September 2010. Demographic factors were associated with chance of responding to the questionnaires.

Interventions = ASQ and MCHAT screenings. ASQ and M-CHAT were returned by 529 (59.3%) of the parents. Parents of white children (390/575 [67.8%]) and those with private insurance (457/660 [69.2%]) were significantly more likely to return screening questionnaires than parents of non-white (64/171 [37.4%]; P < .001) and government-insured children (58/169 [34.3%]; P < .001), respectively. Of the 529 children who returned questionnaires, 109 (20.6%) did not pass at least 1 of the initial screens and 12 (2.3%) were referred after not passing the follow-up screening. ASQ and MCHAT screening by mail is not a sufficient method to comprehensively screen a general population. A nurse-completed, standardized follow-up process after an initially failed ASQ-3 or MCHAT screen increases the yield of appropriate referrals to EI agencies, medical sub-specialists, mental health providers and other professionals.
Marks & Hix-Small, 2009. Medical setting, primary care clinics (general population) (+)

Prospective pre-post observational cohort design. Study involved 64 lower-risk, mostly late-preterm and 1363 term children who originally presented at their 12- or 24-month well-visit. ASQ and Pediatric Developmental Impression (PDI) results were compared. Medical record and county early intervention/early childhood special education (EI/ECSE) follow-up outcomes were conducted at 36 to 60 months. See Hix-Small, Marks et al. 2007 study for other details about the methods and population. Of note, the state Oregon has stricter criteria (requiring the child to be more significantly delayed or possess a high risk biological and/or environmental condition) for EI/ECSE eligibility compared to many other US states.

Mathews, 2014. Medical setting, pediatric developmental subspecialty clinic (at risk population) (+)

Retrospective observational design. Objective: to describe an intake process and results of screening for developmental and autism spectrum disorders in children referred to a tertiary developmental disability center. Methods: A total of 379 referred children younger than 6 years were "prescreened" with the ASQ-3 and the Modified Checklist for Autism in Toddlers (MCHAT) or the Social Communication Questionnaire (SCQ). Medical records were reviewed to identify demographic variables and parental primary concerns. Study was conducted at a metropolitan Midwestern university developmental disabilities center. 2.0% of participants were < 12 months; 12.6% were aged 12 to 24 months; 21.4% were aged 24 to 36 months; 32.4% were aged 36 to 48 months; 24.6% were aged 48 to 60 months; and 7.0% were aged 60 to 72 months. 75.9% were boys. After ASQ-3, MCHAT and SCQ screening, no reference/gold standard or diagnostic testing was done. Study focused on intake procedures at pediatric developmental subspecialty clinics.

McCrae, 2011. Cross-sector setting, (at risk population) (+)

Retrospective observational design: Objectives: (1) What are rates of developmental concern using the ASQ in child welfare? (2) Do results support policies that limit screening to children with substantiated maltreatment or who are placed out-of-home? And (3) how comparable are rates of concerns when screening is conducted by CW and EI providers? Methods: Children ages 0–3 with child welfare (CW) substantiated maltreatment (i.e., childhood neglect and/or abuse) in Pennsylvania were screened with the ASQ and ASQ:SE. Study conducted on over 500 children to address whether children's substantiation status, living situation, and administering worker as CW or EI predicted their ASQ and ASQ:SE screening rates. Bivariate and logistic regression analyses were used. Study uses data from a statewide database of screening (including ASQ and ASQ:SE) results, along with case and family characteristics populated by county CW agencies, and from phone interviews conducted in June and July, 2009. Child welfare representatives from all 67 counties were interviewed by phone during an earlier portion of this study that addressed screening policy and practices. Respondents were asked to report which agency – CW or EI – typically conducts screening in their community.

Morelli, 2014. Mixed methods design. Objective: to identify challenges to developmental screening and strategies for screening in an urban pediatric setting. Methods: Parents of young children and clinicians at 4 urban pediatric practices participated in focus groups prior to the implementation of ASQ-2 and the Modified Checklist for Toddlers (MCHAT). Participants were queried regarding attitudes, social norms, and barriers to screening. Using information from the focus groups, workflow strategies were developed for implementing screening. Referral rates and satisfaction with screening were gathered at the conclusion. Children were eligible to participate if they were ≥ 30 months old at the time of the initial interview, had no major congenital anomalies or genetic syndromes, were never placed in out-of-home foster care, and were not receiving EI services. Clinicians were eligible for participation if they were attending pediatricians, nurse practitioners, or pediatric residents. Study was conducted from December 2008 to June 2010 in the Philadelphia metropolitan area. Parents had positive experiences with the process of screening and did not report any harm.

Intervention = ASQ screenings. At 12 and 24 months, preterm (versus term) referral rates were 9.5% (versus 5.6%) with PDI blind to ASQ results and 26.2% (versus 8.1%) with the ASQ. By 36 to 60 months, 37.5% of preterm (20.8% term) children were referred to an EI/ECSE agency; of which, 50.0% of preterm (42.4% term) children were eligible for services, 54.2% of preterm children were identified with a developmental-behavioral disorder and 29.2% of preterm (20.8% term) children did not follow-up. For ASQ-only preterm referrals, 55.6% were subsequently diagnosed with a developmental delay and/or disorder. Preterm children were ~2 times more likely to be EI/ECSE-eligible than term children. 2-3 years after the initial EI/ECSE referral, 57% of referred children were deemed EI/ECSE-ineligible. Combined referral, quality improvement and outcome data suggests clinicians should lower threshold for using the ASQ for ex-preemies. Diligent surveillance and a more collaborative, standardized, reliable and interpersonal referral process (with care coordination) was recommended.
Nguyen, 2014. Medical setting, primary care clinic (general population) (+) Prospective observational design. To determine if parental deployment affects the cognitive, social and emotional development of preschool aged children in military families. Methods: Demographic & age-appropriate ASQ-3 & ASQ:SE results were collected on children in military families who had a parent who was or was not deployed in order to determine the effects to the cognitive and social-emotional development of preschool age children. 151 eligible children from Fort Bragg, North Carolina. 95 children had a parent who had been deployed during their lifetime. There were no statistically significant differences between the socio-demographics (ethnicity, enlisted rank or anxious/depressed status) of the deployed vs. non-deployed groups. Intervention = ASQ and ASQ:SE screenings. There was a significant difference in ASQ-3 failure rates for children in the deployed group compared to the non-deployed group. Children of deployed parents were at least 2x as likely to have a failed ASQ or ASQ:SE. 30.5% of children who had a deployed parent failed the ASQ or ASQ:SE while 12.5% of children who did not have a deployed parent failed the ASQ or ASQ:SE (P<0.009). On the ASQ:SE, 16.8% of children who had a deployed parent failed vs. 5.4% who did not have a deployed parent (P=0.31). Adverse outcomes could possibly be prevented by the early detection of delays and by more frequent screening and, when concerns arise, care coordination for EI and other community services.

Pizur-Barnekow, 2010, Medical setting, primary care clinics (general population) (+/-) Descriptive and cross-sectional survey. Objective: to describe the range of early identification services available during early childhood. These services include surveillance, screening, and developmental and medical diagnostic evaluation. Methods: Nonprobability sample of approximately 600 healthcare providers in a Midwestern state (Wisconsin). 129 primary care providers participated but only 78 (60%) completed the entire survey. Professionals were specifically recruited if they refer to EI programs, Women, Infants and Children (WIC), Headstart preschools, public health, and childcare. Participants recruited from electronic list serves, via professional newsletters, and at professional meetings so selection bias was a concern. Intervention = survey. The ASQ was found to be the most commonly used standardized developmental screening instrument. The top two barriers to doing standardized developmental screening included: 1) lack of reimbursement 2) family instability. Medical providers also thought that more education was needed about each community’s resources. To avoid duplication of screening, statewide database systems that maintain screening records, similar to those utilized to maintain records of immunizations, may be helpful. This type of research is potentially valuable but selection bias was a concern.

Porte, 2016, Medical setting, primary care clinics (general population) (+/-) Descriptive and cross-sectional survey. Objective: to investigate current developmental surveillance and developmental screening practices by pediatric primary care providers in a diverse New Jersey county. Methods: Survey used to investigate current developmental surveillance and developmental screening practices by primary care providers in a diverse New Jersey county. 217 providers were contacted with a final sample size of 57 pediatric primary care respondents (26% response rate despite being offered incentives) from 13 different municipalities. It was estimated that 52.4% of the children in this community are African American/Black and 33.8% are Hispanic/Latino. Intervention = survey. Medical providers also thought that more education was needed about each community’s resources. To avoid duplication of screening, statewide database systems that maintain screening records, similar to those utilized to maintain records of immunizations, may be helpful. This type of research is potentially valuable but selection bias was a concern.

Quigg, 2013. Medical setting, pediatric oncology (at risk population) (+) Prospective observational pilot study. Objectives: to understand whether implementation of a protocol for developmental screening could identify early effects of chemotherapy on child development, and if systematic screening would be feasible and allow for initiation of appropriate supportive care services among young cancer patients. Methods: ASQ-3 used to screen young oncology patients who were 4 to 48 months of age with newly diagnosed cancer. Subjects screened within 28 days of diagnosis (baseline), and at 6 and 12 months (after beginning cancer treatment). Exclusion criteria were having a benign hematology diagnosis or non-English-speaking parents. Setting = children’s cancer center in Indiana. Intervention = ASQ screening. 26 of 30 enrolled parents (87%) completed all 3 screens. ASQ completed by parents in 15 minutes. ASQ-3 identified unsuspected developmental delays as follows: 7 at baseline, 4 at 6 months, and 3 at 12 months (after beginning cancer treatment by standard protocols). ASQ-3 was feasible, identified unsuspected delays in young children with cancer, and helped to initiate appropriate referrals (eg, early intervention agencies and developmental sub-specialties). No reference/gold standard or diagnostic neuropsychological battery of testing was done.

Radecki, 2011. Medical setting, primary care clinics across the USA (general population) (+) Descriptive cross-sectional survey. Objective: to compare pediatricians’ use of standardized developmental screening tools from 2002 to 2009. Methods: a national, random sample of non-retired US members of the American Academy of Pediatrics were mailed Periodic Surveys (2002: N=1617, response rate 55%; 2009: N=1620, response rate was 57% so selection bias was a possibility). Chi-squared analyses were used to examine responses across survey years; a multivariate logistic regression model was developed to compare differences in using ≥1 formal screening tools across survey years while controlling for various individual and practice characteristics. Intervention = survey. The percentage of providers who self-reported “always/almost always” using ≥1 standardized developmental screening tools more than doubled from 2002 to 2009 (from 23.0% to 47.7%), as did use of specific instruments like the ASQ. No differences were noted on the basis of physician or practice characteristics. Additional research is needed to identify barriers to improve the use of standardized developmental screening tools (like the ASQ) in US primary care medical practices.

Roane, 2012. Medical setting, primary care clinics (general population) (+/-) Descriptive cross-sectional survey. Objective: to evaluate whether positive/concerning ASQ results would prompt physicians to refer for additional evaluation or EI as recommended by the American Academy of Pediatrics algorithm. Methods: A random sample of 207 physicians (from 15 states) read one of three hypothetical clinical vignettes describing an 18-month-old child with ambiguous language development. Vignettes differed on the presence or absence of an ASQ score and, if present, on whether the ASQ score was positive or negative/typical. Physicians indicated whether they would refer for further evaluation, EI or other interventions. Intervention = survey. Multinomial regression analyses showed physicians referred more often for further evaluation or EI if the hypothetical ASQ was positive. Likewise, physicians referred less often if the ASQ was negative. Physicians without the ASQ scores did not choose one action more frequently over another. Physicians do refer, as recommended, when presented with positive ASQ results. Given the use of hypothetical vignettes, it’s important to investigate whether positive ASQ results impact referrals in primary care practices. Bias concerns related to the theoretical nature of the vignettes.
Ryan-Krause, 2009. High school childcare center (at risk population) (+) Descriptive observational design. Objective: to compare adolescent mothers’ subjective perceptions of their children’s development with objective developmental assessments. Methods: Volunteer sample of 45 adolescent mother/child pairs recruited from an urban high school childcare center in New England. 33 mothers (mean age = 17.4 years, mean grade level = 11th grade) completed the ASQ-2 and the Bayley Scales of Infant Development (BSID) was administered to their children of adolescent mothers with the mothers being present. Nearly 71% of adolescent mothers were in the appropriate grade and 4% attended special education classes. 55% African American, 26% Latina (26%), 3% white, and 16% diverse ethnicity.

San Antonio, 2014. Medical setting, primary care clinics (general population) (+) Prospective randomized design (with control group being ASQ administration under non-standardized conditions). Objective: to examine the reproducibility of ASQ examined under standardized vs non-standardized conditions in an underserved population because developmental screens are often used in non-standardized conditions, such as pediatric waiting rooms, despite typically being validated under standardized conditions. Methods: English- or Spanish-speaking parents of 18- or 30-month-old children completed the ASQ-3 in the waiting room and then were randomized to repeat the ASQ in waiting room (W-W) or with standardized (W-S) conditions. ASQ-3 fail rates and intra-class correlation coefficient for each of 5 ASQ domains was calculated. 131 parents were randomized (66 W-W, 65 W-S).

Schatz, 2017. Medical setting, pediatric hematology subspecialty clinic (at risk population) (+) Prospective observational design. Objectives: 1) to examine syntactic processing scores in higher-risk SCD genotypes and compare them to lower-risk genotypes. 2) to examine if a targeted screening approach would show better sensitivity to genotype risk than a parent-report broad-band developmental screening tool. Methods: 74 4-year-old children with sickle cell disease (SCD) were screened using the ASQ-2 and Fluharty Preschool Speech and Language Screenings Test, 2nd edition to assess if biomedical risk factors for neurologic disease are related to developmental screening outcomes in preschool aged children. Genotype and other biomedical variables were coded from medical records. 100% of the children were African-American.

Shaw & Hatton, 2009. Medical setting. EI agencies (at risk population) (-) Cross-sectional survey. Objective: to query state Part C and Section 619 coordinators regarding screening measures, diagnostic instruments and procedures, and trends in identifying young children with autism spectrum disorders ≤ 5 years of age. Methods: 2008, online survey obtained from 40 respondents in 30 states/jurisdictions throughout the US and Pacific jurisdictions. 18 were Part C (EI) program coordinators, 13 were Section 619 program coordinators, and 9 indicated that they represented both programs.

Sheldrick, 2012. Medical setting, primary care clinics (general population) (+/-) Retrospective observational design. Objective: To describe responses to the questions “Do you have concerns about your child’s behavior? Development? Learning?” among parents seeking pediatric care, and to analyze their correspondence to the ASQ-3 and ASQ-SE. Study described responses to the questions “Do you have concerns about your child’s behavior? Development? Learning?” among parents seeking pediatric care, and to analyze their correspondence to the ASQ-3 and ASQ-SE. Of 465 parents of children aged 3 to 65 months recruited at pediatric primary care practices in Greater Boston, 465 provided complete data for analysis. After completing a questionnaire that asked whether they had any or all of these concerns, parents filled out a developmental screener (ASQ-3) and a behavioral screener (ASQ-SE); however, there was no reference standard/diagnostic testing so the clinical applicability of this study is questionable. Parents from 4 urban (n=169) and 3 suburban (n=282) primary care pediatric practices. Only 62% of the parents approached enrolled and returned the ASQ-3 and ASQ-SE questionnaires. This study might have had a conflict of interest issue. Study authors are authors of the Survey of Well-Being of Young Children (SWYC), which uses the studied questions about parental concerns.

Intervention = ASQ and BSID. On the BSID, group mean scores all fell in normal range. However, almost 20% of children had ≥1 delays. Almost 73% of teen mothers accurately assessed their children’s development on the ASQ-2 when compared with the BSID. 18% of mothers suspected delays when no delays were identified on the BSID. A single mother identified delay in a different domain than that identified on the BSID. Findings support 1) children at risk for delays, (2) although teen mothers varied in their abilities to assess their children’s development, these teen mothers benefited from completing the ASQ-2.

Intervention = ASQ screenings. In an underserved population, 25.8% failed the first ASQ-3 screen completed in the waiting room before randomization. There was no significant difference in fail rates between study arms on the first or second screen. Intra-class correlation coefficient for W-W in the 5 domains ranged from 0.66 to 0.95, and for W-S from 0.73 to 0.92. There were no significant differences between intra-class correlation coefficients in W-W versus W-S in any domain. The ASQ, when completed in the waiting room, is reliable compared with standardized conditions. It’s okay to administer the ASQ in waiting rooms. Unclear if physicians interpreting the ASQ-3 were blinded to W-W vs. W-S groups.

Intervention = ASQ and language screening. 52 children with higher-risk SCD genotypes showed lower performance than 25 children with lower-risk genotypes on measures related to neurologic disease risk in older children (syntactic processing). Genotype risk was related to rates of a positive ASQ-2 (52% positive screenings in high-risk genotypes vs 12% in low-risk genotypes). Screening results were related to transcranial Doppler ultrasound results assessing cerebral blood flow. One conclusion was that the 4-year interval ASQ helped to identify SCD’s neurodevelopmental effects.

Intervention = survey. 83% of EI programs screen for autism using the ASQ:SE and 73% use the Modified Checklist for Autism in Toddlers (MCHAT); however, the majority use a wide variety screening tools and diagnostic measures to identify autism. Further guidance is needed to help states use the most evidence-based tools and strategies to identify autism. Online survey methodology increases the risk of selection bias. Other methodological flaw was that the survey response rate was not reported.

Intervention = surveillance questions about parental concerns + ASQ-3 and ASQ:SE screenings. 108 parents (24%) reported having at least 1 concern about their child. Greater child age, male gender, and lower family income were associated with more concerns about development, behavior, and learning. Moderate agreement was found between parents’ concerns and their responses on screening instruments, but among parents who identified no concerns, 18% were identified as at risk on one or both screeners. Compared with children who were not identified on either screen, parents of children identified only on the ASQ:SE were more likely to have behavior concerns. Parents of children identified on ASQ and ASQ:SE were more likely to have noted concerns about both behavior and development. No specific type of concern was associated with identification on the ASQ-3 alone. Parents’ self-report of concerns showed moderate agreement with the results of ASQ-3 and ASQ:SE. Agreement was higher for behavioral than developmental concerns. Authors are the authors of the SWYC screening tool so conflict of interest may have been an issue.
<table>
<thead>
<tr>
<th>Sikes, 2009.</th>
<th>Prospective observational design. Objective: to describe the agreement between the ASQ and the Parents Evaluation of Developmental Status (PEDS) in a pediatric primary care setting. Parents/caregivers of 60 children (ages 9 to 31 months) completed the PEDS and ASQ-2 at the same well-visit. Concordance (PEDS and ASQ-2 results agree) and discordance (results differ) for the 2 screens were determined. Mean age of the children was 17.6 months, 77% were Hispanic, and 50% of their parents had a high school education or less. Secondary participants were 6 primary care pediatricians from an academically affiliated community hospital–based practice in Northeast Ohio. No reference/gold standard testing was done to see which tool, the ASQ or PEDS, had better accuracy.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squires, 2002.</td>
<td>Prospective observational design. Objective: to describe and summarize the procedures and outcomes of the Oregon Healthy Start developmental screening project including the ASQ and its psychometric properties as well as the Oregon Healthy Start Program. Recommendations are given for use of parent-completed measures like the ASQ in home visit/early learning programs. Methods: All “Healthy Start” home visitors/early childhood educators received training on the use of the ASQ. Parents completed the ASQ when children were 4, 8, 12, 18, 24, 30, 36, and 48 months of age. Children and families resided in 16 Oregon counties. Under half the mothers had completed high school or received a GED (47%). 72% of parents were White and 24% were Hispanic or Latina.</td>
</tr>
<tr>
<td>Wang, 2013.</td>
<td>Observational (unclear if prospective or retrospective) design. Objective: “to examine the well-rounded early childhood development in a county as large as the state of New Jersey, a comprehensive Context, Input, Process, and Product model was employed to determine pertinent variables from the current research literature and reconfirm the relationship through empirical data analyses.” Main text also listed 3 more detailed “research objectives”. Methods: Empirical data were gathered from 131 children at age 3 years to assess coexisting effects from area-specific services at 21 preschools across Kern County, California.</td>
</tr>
<tr>
<td>Welsh, 2012.</td>
<td>Prospective observational design. Objective: to identify factors related to developmental recovery in a diverse sample of internationally adopted children. Methods: This study followed 106 international adoptees over a 18-month period. Adopting mothers completed the ASQ-2 (even though the ASQ-3 was published in 2009) at 6, 12, and 24 months post-adoption, assessing their children’s development in multiple domains. No reference standard/gold or diagnostic testing was done. Another study limitation is that an 18-month follow up period might not have been long enough for this at-risk population.</td>
</tr>
<tr>
<td>Wiley &amp; Meinzen-Derr, 2013.</td>
<td>Prospective observational design. Objective: to detect additional developmental delays with ASQ in children already diagnosed with bilateral sensorineural hearing loss. Methods: Participants included 50 children with any degree of bilateral sensorineural hearing loss, ages 0 to 6 months at the time of recruitment and 0 to 6 months of age at the time of the second test. Participants were recruited from an urban tertiary referral center at a children’s hospital and they were recruited through EI programs serving children who are deaf/hard of hearing. Recruitment strategy is realistic and practical but may have increased the risk of bias. Note: Children with hearing loss are automatically EI eligible in the state of Ohio and the large majority of US states.</td>
</tr>
<tr>
<td>Windham, 2014.</td>
<td>Prospective observational design. Objective: to detect developmental delays and autism in an underserved, primarily Hispanic population. Methods: Study involved 1,760 children from a typically under-served, primarily Hispanic, population. Between January 2008 and July 2009, routine screening was implemented in children 16–30 months of age at well-visits using the ASQ-3 and Modified Checklist for Autism in Toddlers (M-CHAT). Parents had to be English or Spanish-speaking in order to complete the screenings. All eligible children who returned to the clinic for a non-urgent appointment during the study period were eligible for being re-screened, regardless of their scores on the first ASQ-3 or MCHAT screening.</td>
</tr>
</tbody>
</table>

**Intervention = ASQ and PEDS screenings.** Overall, 37% failed the PEDS; 27% failed the ASQ. 31 children passed (52%) both screens; 9 (15%) failed both; and 20 (33%) failed 1 but not the other (13 PEDS and 7 ASQ). Agreement between the 2 screens was only fair, statistically no different from agreement by chance. There was substantial discordance between PEDS and ASQ-2. Although these are preliminary data, clinicians should be aware that when implementing AAP recommendations, the choice of screening instrument may affect which children are likely to be identified for additional evaluation. Despite lack of reference/gold standard testing, this study did answer its stated objectives and is impactful.

**Intervention = ASQ screenings.** Results indicate the ASQ 1) structured home visits, 2) engaged parents in viewing their children in a strengths-based perspective (by asking parents to evaluate what activities their children were currently doing), 3) allowed low-income parents to complete a screening with little trouble, 4) provided concrete suggestions for early learning (geared to the child’s current developmental level), 5) offered a parent-centered and structured way to observe and comment on their child’s development and behavior. Further research needed regarding use of ASQ in “Healthy Start” program. Study has real-world implications but the lead author of this study is also the lead author of the ASQ so there is a potential conflict of interest.

**Intervention = ASQ screenings.** Sample overall demonstrated linear improvement over time in most ASQ domains, but internationally adopted children with initially low ASQ scores remained significantly lower than others at the 18-month follow-up. ASQ scores were unrelated to adoption age, but significant differences by birth country emerged. Across most domains, children born in Eastern Europe had lower scores than children born in other regions. Main implication is that the ASQ was successfully used to monitor developmental milestones in international adoptees as they grew older.

**Intervention = ASQ screening and the Revised Gesell Developmental Schedules.** While 32% had a delay outside the communication domain, the ASQ had poor sensitivity on the overall score without the communication domain, as well as for fine motor, problem-solving, personal-social domains. ASQ had good sensitivity for the communication and gross motor domains and good specificity (ranging 83–85%) on specific domains as well as for the overall score (70%). ASQ does not effectively identify additional domain delays (beyond the communication domain) for children already identified with bilateral sensorineural hearing loss.

**Intervention = ASQ and MCHAT screenings.** Screen positive/concerning rates of 39% and 26%, respectively, were higher than previous reports. Hispanics were more likely to score MCHAT positive than non-Hispanics (adjusted OR 1.3, 95% CI 1.2–2.4), as were those screened in Spanish. About 30% of screen-positive children were referred for further assessment, but only half followed up. ASQ-3 and MCHAT screening in this population is feasible, but may require additional resources and care coordination. Attention to the cultural applicability of screening instruments, as well as to explaining the results or need for additional services to parents, is critical to serve the growing Hispanic population.
| Yeung, 2017.* | Prospective cohort study. Upstate KIDS recruited 5034 mothers from New York State (excluding New York City) at ~4 months postpartum between 2008 to 2010. Parents completed the ASQ-2 when their children were 4, 8, 12, 18, 24, 30, and 36 months of age corrected for gestation. Analyses included 3759 singletons and 1062 nonrelated twins with ≥1 ASQ-2 screening questionnaires returned. Adjusted odds ratios (aORs) and 95% confidence intervals were estimated by using generalized linear mixed models accounting for maternal covariates (age, race, education, insurance, marital status, parity, and pregnancy smoking). Of note, about 20-30% of US adults are obese with a body mass index (BMI) > 95th percentile. About 20% of US women enter into pregnancy with obesity (a BMI ≥30). | Intervention = ASQ screenings. Compared with normal/underweight mothers (BMI <25), children of obese mothers (26% with BMI ≥30) had increased odds of failing the ASQ-2 fine motor domain (aOR 1.67; confidence interval 1.12–2.47). The association remained after additional adjustment for paternal BMI (1.67; 1.11–2.52). Paternal obesity (29%) was associated with increased risk of failing the ASQ-2 personal-social domain (1.75; 1.13–2.71), albeit attenuated after adjustment for maternal obesity (aOR 1.71; 1.08–2.70). Children whose parents both had BMI ≥35 were likely to additionally fail the ASQ-2 problem-solving domain (2.93; 1.09–7.85). Findings suggest that maternal and paternal obesity are each associated with specific ASQ-domain delays in early childhood, emphasizing the importance of family information (like a family history of maternal or paternal obesity) when screening for developmental delays as recommended by American Academy of Pediatrics (AAP). |
| Yovanoff, 2013. | Prospective observational design. When differences due to mode of completion (i.e., paper–pencil, computer-based) are present, threats to measurement validity are posed. Administration mode effects (of web-based and paper–pencil versions of the ASQ-2) were estimated. Setting this study apart from similar studies reported in the literature, the ASQ-2 requires parents to observe and rate their children’s skills. Recruitment procedures for traditional pencil-paper ASQ-2 included solicitation of parents by early childhood and Early Head Start providers across the United States. ASQ study was posted on early childhood sites such as daycareresource.com and craigslist.com. Interested caregivers visited the ASQ website by using various search words (e.g., parent help, play activities, stages of development, child research, parent education, home school, child progress) which could have increased selection bias. | Intervention = ASQ (paper-pencil vs. online) screening. Using item response model invariance testing procedures, analyses tested whether ASQ-2 items administered via the Internet function differently from corresponding traditional paper–pencil items. Analyzing the 4-, 12-, and 24-month ASQ-2 intervals, statistically significant differences (DIF or differential item functioning) were obtained on 10 of the 90 items examined in this study. Although DIF was observed for some items, the overall DIF model was rejected for all domains. On the basis of these results, the paper–pencil and web-based measures can be considered equivalent and the mode effect is not present; ASQ measures obtained from either mode are therefore interchangeable. Noteworthy remedies are considered for web-based administration of the 10 specific items lacking invariance. Implications of this study are applicable to clinic settings. |
Table 3. Scandinavian studies where ASQ and/or ASQ:SE screening was universally performed in a general or at-risk population. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.

<table>
<thead>
<tr>
<th>Study identification &amp; setting</th>
<th>Overall level of evidence</th>
<th>Comments about study design, objectives, methods (including population)</th>
<th>CASP key question = “Are the results of the study valid?”</th>
<th>Comments about interventions and outcomes (results and conclusions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alvik, 2014. Cross-sector setting from birth registry in Norway (general population)</td>
<td>(+)</td>
<td>Longitudinal prospective cohort design. Objective: to explore variables predicting low developmental scores in 6-month-old infants in a population-based study. Methods: Women from Oslo, Norway completed questionnaires at 17 and 30 weeks of pregnancy and 6 months post-term; N=1053 after exclusions (women with non-Scandinavian ethnicity, twin births, infants &lt;3.5 or &gt;7.0 months corrected age, and birth weight &lt;2.5 kg). Data collected from Norwegian Birth registry. Measures included socio-demographic variables, maternal mental health and pregnancy life style, data concerning the birth/newborn, and the 6-month interval ASQ-2. Logistic regression analyses used to identify variables predicting an ASQ-2 score ≤15th percentile.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
</tr>
<tr>
<td>Janson 2003*. Categorized as a psychometric study in Norway but added to this table because of its relevance for universal screening in Scandinavia.</td>
<td>(+)</td>
<td>Nationwide/cohort norming study. Objective: To describe the influence of demographic variables on participation rate in ASQ-2 completion, and to discuss the implications for data analysis and design of future similar studies. Methods: Data from a nationwide/cohort sample for the primary purpose of norming the Norwegian ASQ-2 using 2392 mother-child dyads. Twins not included. 240 mothers completed the ASQ at 3 months, and 120 mothers (for each age-interval) of children aged 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 54 and 60 mo. The bivariate and multivariate influence of demographic variables on responding was investigated to describe the influence of demographic variables on ASQ-2 participation rate, and answer other study objectives.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
</tr>
<tr>
<td>Junge, 2017. Cross-sector setting from birth registry in Norway (at-risk population)</td>
<td>(+)</td>
<td>Prospective longitudinal cohort design. Objective: to investigate if maternal depression at different time points during the perinatal period impacts children’s social- emotional development at 2 years of age. Methods: Participants = 1255 women who gave birth at Akershus University Hospital in Norway(Akerhus Birth Cohort or ABC). Maternal depressive symptoms assessed by using the Edinburgh Postnatal Depression Scale (EPDS) at pregnancy week 32 and at 8 weeks and 2 years postpartum. Children’s social-emotional development was assessed with ASQ:SE at age 2 years. Only 27% of the initial cohort’s participants followed up at 2 years. Bi- and multivariate logistic regression analyses was conducted to examine links between maternal peripartum depression and children’s early social-emotional development.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
</tr>
<tr>
<td>Moe, 2016. Cross-sector setting from birth registry in Norway (at-risk population).</td>
<td>(+)</td>
<td>Prospective longitudinal cohort design. Objective: to investigate early infant social withdrawal, maternal symptoms of depression and the child’s social emotional functioning at 12 months. Methods: The sample consisted of 248 of full-term infants and their mothers, and a group of 64 moderately premature infants and their mothers. At 3 months, the infants were observed with the Alarm Distress Baby Scale (ADDB) and the mothers completed the Edinburgh Postnatal Depression Scale (EPDS). At 12 months, the mothers filled out questionnaires about the infants’ social emotional functioning (Infant Toddler Social Emotional Assessment and the Norwegian ASQ:SE, 1st edition – which was published in 2002 or 14 years before this study was published). ASQ:SE-2 was published in 2015.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
<td>Intervention = ASQ screenings. No systematic reviews with or without meta-analysis, or randomized controlled trials (with lower potential for bias) were identified. No qualitative survey studies (with higher potential for bias) were identified. Study appraisal was done with the Critical Appraisal Skills Programme (CASP) Checklists for systematic reviews, randomized controlled trials, cohort studies, case control studies, diagnostic test studies, qualitative research and clinical prediction rule studies. The CASP has 10 to 12 items, which considers the clarity of the research aims, appropriateness of methods, subjects, data collection and the value of the study. A specific critical appraisal score is not generated from using the CASP Checklists; however, they were used to help assign an overall “level of evidence” rating to each study according to those set by the Dutch Institute for Healthcare Improvement. (+) = Credible study with low-to-medium bias. (+/-) = Dubious credibility and medium-to-high or unclear bias. (-) = Minimal credibility and high bias.</td>
</tr>
</tbody>
</table>
Prospective observational design with randomized sample. Objective: to obtain ASQ scores from the background population so that these may be used as a reference group for extremely preterm children at nine and 18 months of corrected age. Methods: 298 children randomly chosen among general Danish population in three different groups: 9-, 18- and 21-month-old children. In contrast to most other ASQ studies, the authors purposely did not administer the correct age-interval ASQ to parents. Parents used ASQ intervals that were designed for older children (for the purpose of creating a reference group for extremely preterm children at 9 and 18 months of corrected age). Authors stated this was done “to allow the best functioning children to be appropriately represented.” The parents received the 10-month ASQ-2 when the child had reached a corrected age of 9 months and the 24-month ASQ-2 when their child had reached the corrected age of 18 or 21 months. Histograms were made for ASQ total score. Mean, median and standard deviation calculated. Linear regressions performed on ASQ total score and adjusted age to evaluate association between age and ASQ score.

Cross-sector but also medical setting in Norway (general population)

Population-based prospective cohort study. Objectives: 1) to determine typical developmental pathways in infancy within communication, gross motor, fine motor, problem-solving and personal-social skills from 4 to 24 months; 2) to examine the degree to which these pathways are predicted by sex, gestational age, Apgar score ≥ 7, mother’s ethnic origin, mother’s educational level and maternal depressive symptoms at 6 weeks post partum. Methods: Study done because there was limited epidemiological research (with diverse results) describing population-based samples regarding developmental pathways throughout infancy. Identifying predictors for developmental pathways can inform early intervention services. The ASQ-2 was used to measure communication, gross motor, fine motor, problem-solving and personal-social skills longitudinally in a large, population-based sample of 1555 infants recruited from well-baby clinics in 5 municipalities in southeast Norway. Fairly low attrition rates. Latent class analyses used to identify common pathways within the five developmental areas.

Cross-sector but also medical setting in Norway (general population)

Population-based prospective cohort study. Objective: to estimate prevalence rates of SDD among infants at 4, 6 and 12 months of age based on parent-completed ASQ, and to investigate associations of suspected developmental delays (SDD) with gender, gestational age < 37 weeks and maternal education. Methods: Prevalence estimates on SDD in young infants are scarce and a necessary first step for planning EI services. Prevalence of SDD estimated at 4, 6 and 12 months, in addition to associations of SDD with gender, prematurity and maternal education. Study based on a Norwegian longitudinal sample of 1555 infants and their parents attending well-baby clinics for regular health check-ups. Parents completed the Norwegian ASQ-2 prior to check-ups, with a corrected gestational age being used to determine the time of administration for preterm infants. Scores ≤ the established cut-offs in one or more of the five development areas: communication, gross motor, fine motor, problem solving and personal-social, which defined SDD for an infant were reported. Chi-square tests performed for associations between the selected factors and SDD.
<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td>Page 1</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
<td>Page 2</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td>Pages 3 to 8</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>Page 3</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
<td>Not available</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
<td>Pages 8 &amp; 9. Table 1</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
<td>Pages 8 &amp; 9</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td>Page 8. Figure 1</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td>Pages 8 &amp; 9. Table 1</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td>Page 9</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
<td>Page 9</td>
</tr>
</tbody>
</table>
# PRISMA 2009 Checklist

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
<td>Page 9</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
<td>Page 9</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I²) for each meta-analysis.</td>
<td>Page 9</td>
</tr>
</tbody>
</table>

## RESULTS

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study selection</td>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
<td>Pages 9 &amp; 10. Figure 1.</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
<td>Pages 9-16. Tables 3 &amp; 4.</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).</td>
<td>Pages 9 to 16. Tables 3, 4 &amp; 5.</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
<td>Outcomes on Tables 3 &amp; 4. Article categorization</td>
</tr>
<tr>
<td>Item</td>
<td>Description</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Synthesis of results</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency. Meta-analysis not done.</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Risk of bias across studies</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15). Tables 3, 4 &amp; 5.</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Additional analysis</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). Figures 2 &amp; 3. Meta-analysis not done.</td>
<td></td>
</tr>
</tbody>
</table>

**DISCUSSION**

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Summary of evidence</td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). Pages 16-22</td>
</tr>
<tr>
<td>25</td>
<td>Limitations</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). Page 19</td>
</tr>
<tr>
<td>26</td>
<td>Conclusions</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research. Pages 19-22</td>
</tr>
</tbody>
</table>

**FUNDING**

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Funding</td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. Non-funded study. No financial disclosures or conflicts of interest.</td>
</tr>
</tbody>
</table>


For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).
PRISMA 2009 Checklist