Evidence Synthesis

Number 129

Screening for Autism Spectrum Disorder in Young Children: A Systematic Evidence Review for the U.S. Preventive Services Task Force

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2012-0001-5I

Prepared by:
Vanderbilt Evidence-based Practice Center
Vanderbilt University Medical Center
Nashville, TN

Investigators:
Melissa L. McPheeters, PhD, MPH
Amy Weitlauf, PhD
Alison Vehorn, MS
Cora Taylor, PhD
Nila A. Sathe, MA, MLIS
Shanthi Krishnaswami, MBBS, MPH
Chris Fonnesbeck, PhD
Zachary E. Warren, PhD

AHRQ Publication No. 13-05185-EF-1
August 2015

This report is based on research conducted by the Vanderbilt Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2012-0001-5I). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients).

The final report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Acknowledgments

The authors acknowledge the following individuals for their contributions to this project: Dr. Evelyn Whitlock, Ms. Tracy Biel, Dr. Mary Louise Lindegren, Ms. Sanura Latham, Ms. Rebecca Jerome, and Ms. Jessica Kimber. The authors thank AHRQ Officers Elizabeth Kato, MD, MRP, Karen Lee, MD, MPH, Tess Miller, DrPH, and Tracy Wolff, MD, MPH, as well as current and former members of the U.S. Preventive Services Task Force who contributed to topic deliberations.

Structured Abstract

Context: The U.S. Centers for Disease Control and Prevention estimates that one in 68 children has an autism spectrum disorder (ASD) and the majority of children are not diagnosed until after 4 years of age. Current approaches rely on developmental surveillance, general developmental screening, and/or parental concerns. Systematic screening has been advocated for identifying ASD at earlier ages.

Objective: We systematically reviewed the evidence about benefits and harms of routine screening for ASD in primary care settings.

Methods: We explicitly focused on studies of screening instruments for use in young (≤36 months of age), unselected populations (e.g., universal screening approaches).

Results: We identified 17 unique screening studies reported in 22 papers. The most commonly studied tool was the Modified Checklist for Autism in Toddlers (M-CHAT) including the most recently available variant (M-CHAT-Revised with Follow-Up [M-CHAT-R/F]), which has a positive predictive value of 48 percent in diverse populations of children ages 16 to 30 months. Forty-two studies of good and fair quality addressed interventions for young children. Among these, 17 involved direct provision of intervention to children. Fifteen of these 17 studies assessed cognitive outcomes, and outcomes were significantly more improved in the treatment arm vs. comparison arm in 10. Sixteen of these 17 studies assessed language outcomes, and outcomes were significantly improved in the treatment vs. comparison group in 10 studies. Thirteen studies involved parent training. Five of these thirteen studies addressed cognitive outcomes, and outcomes were significantly improved in the treatment vs. comparison group in one study. Twelve of the 13 studies addressed language outcomes, and outcomes were significantly improved in the treatment vs. comparison group in three studies. Thus, 20 studies overall measured cognitive outcomes and 11 reported greater benefit for the intervention group compared with control groups, and language outcomes were significantly improved in treatment vs. comparison arms in 13 of 28 studies assessing language. Twelve studies focused on play and interaction and typically measured joint attention as the outcome. Nine out of 10 studies evaluating joint attention outcomes reported greater benefit in the treatment arm compared with the control arm. None of the studies focused on screen detected children.

Conclusions: Both the M-CHAT and the M-CHAT-R/F, when including the follow-up interview procedure, have PPVs of around 50 percent in community practices, for children between 16 and 30 months of age. Screening tools are widely available. Multiple treatments are available to young children with ASD. Early intensive interventions demonstrate statistically significant improvements in cognitive and language outcomes in children, compared with eclectic treatments obtained in the community or other comparison groups, although the studies are generally small, and within the studies, some children benefit while others do not. We found no studies that directly compared long-term outcomes of screened versus non-screened children. More research is needed to determine the benefits and harms of screening the general population.

Table of Contents

Figures

Tables

Appendixes

Chapter 1. Introduction

The U.S. Centers for Disease Control and Prevention (CDC) estimates that one in every 68 children has an autism spectrum disorder (ASD). Most are diagnosed at or after age 4.[1] Some children are identified early for evaluation and diagnosis, either because of parental concern or through general developmental surveillance by primary care providers. In addition to this subgroup, the screening population for primary care includes other children whose families and providers do not yet have concerns. An estimated 42 to 55 percent of pediatricians regularly screen for ASD in toddlers,[2] with providers less likely to screen toddlers from underrepresented ethnic and language groups (e.g., 29% of primary care pediatricians report offering Spanish ASD screening in populations of children whose parents are native Spanish speakers).[3] Current approaches that include pediatric surveillance, general developmental screening, and a reliance on parents to raise concerns do not identify most children with ASD prior to age 4. The question of whether it is beneficial to do so is a component of this review.

Condition Definition

ASD is defined in terms of persistent, significant impairments in social interaction and communication as well as restrictive, repetitive behaviors and activities.[4] Social communication and social interaction features include deficits in social-emotional reciprocity (e.g., deficits in joint attention, atypical social approach and response, conversational challenges, reduced sharing of interest, emotions, and affect), deficits in nonverbal communication (e.g., atypical eye contact, reduced gesture use, limited use of facial expressions in social interactions, challenges understanding nonverbal communication), and deficits in forming and maintaining relationships (e.g., diminished peer interest, challenges joining in play, difficulties adjusting behavior to social context). ASD features of restricted, repetitive patterns of behavior, interests, or activities may include stereotyped motor mannerisms, use of objects, or speech (e.g., simple motor stereotypies, repetitive play, echolalia, and formal or idiosyncratic speech); insistence on sameness, inflexible adherence to routines, or ritualized patterns of behavior (e.g., distress at small changes, rigid patterns of thought and behavior, performance of everyday activities in ritualistic manner); intense preoccupation with specific interests (e.g., strong attachment to objects, circumscribed or perseverative topics of interest); and sensory sensitivities or interests (e.g., hyper- or hypo-reactivity to pain and sensory input, sensitivity to noise, visual fascination with objects or movement).[5-7] These symptoms cause impairment across many areas of functioning and are present early in life. However, impairments may not be fully evident until environmental demands exceed children's capacity. They also may be masked by learned compensatory strategies later in life. Many children with ASD also have intellectual impairment or language impairment, and the disorder has been associated with known medical, genetic, or environmental factors.

Previously, disorders considered a part of the autism spectrum were divided into discrete categories including: Autistic Disorder, Asperger's Disorder, and Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS), as defined by the *Diagnostic and Statistical Manual of Mental Disorders, Fourth edition* (DSM-IV).[4] The DSM-5, published in May 2013,

combined the previous categorical disorders into a single category of "Autism Spectrum Disorder," with varying degrees of severity depending on the amount of support required by an individual.[8] Because no medical or biological marker exists for ASD, the diagnosis is behaviorally based. Diagnosis is established with a combination of history, observation, and/or formal testing, which may include ASD-specific screening and assessment instruments.[9,10]

Burden of Disease/Illness

The prevalence of ASD in the United States is estimated at 14.7 cases per 1,000 children, or 1 in 68, with estimates varying widely by region of the country, sex, and race/ethnicity.[1] More males (1 in 42) than females (1 in 189) are affected. For many individuals, core symptoms of ASD improve with intervention and maturation;[11-13] however, early core deficits typically translate into varying developmental effects that remain throughout the lifespan, with differing impact on long-term functional outcomes.[14] Many adults with ASD do not obtain traditional markers of adaptive independence, including jobs, independent living, and educational attainment.[15-19]

Etiology and Risk Factors

ASD has a strong genetic component, with heritability estimated to be between 40 and 90 percent.[20-22] A range of genes is implicated in susceptibility to ASD;[22-24] however, environmental exposures and context also play a role in ASD development and neurogenetic expression.[24,25] Identification of specific genetic risk variants has been challenging, and many researchers suggest that there are be multiple pathways involved, including prenatal and postnatal insult in some cases.[23] Current research[26,27] suggests that certain metabolic and other maternal conditions (such as diabetes, hypertension, obesity, and influenza infection) during pregnancy may be associated with increased risk of ASD in offspring. Other studies have investigated the role of advanced maternal and paternal age,[28-30] intrapregnancy interval,[31,32] pesticide exposure,[33] and exposure to mercury and other heavy metals,[34] among other potential risk factors.

In addition to the potential causative genetic and environmental factors described above, being the sibling of another child diagnosed with ASD triples the risk of receiving an ASD diagnosis from 6.7 to 18.7 percent.[35,36] This risk varies by gender and increases twofold when two or more older siblings have ASD. Increasingly, researchers are attempting to follow infant siblings from very early ages in order to better understand the earliest potential actionable features of the disorder.[5]

Rationale for Screening and Clinical Practice Parameters

A recent CDC report found that most children with ASD in the United States are diagnosed at a median age of 4 years, 5 months.[1] On average, median age of diagnosis is somewhat earlier for children with autistic disorder (4 years) than for children with the more broadly defined autism spectrum diagnoses, such as pervasive developmental disorder-not otherwise specified (4 years, 2 months) and Asperger Syndrome (6 years, 2 months). Substantial racial/ethnic differences in

the age of ASD diagnosis also have been documented. For example, one study reported that African-American children received ASD diagnoses an average of 1.4 years later than white children,[37] and CDC figures suggest that at 8 years of age far fewer Hispanic and African-American children have been identified with ASD.

Supporters of universal screening say that delays in accurate diagnosis may contribute to familial distress[38] and limit access to intervention services,[39,40] which many experts consider to be important for improving children's short-term and longer term outcomes. Although some children can access treatments through early intervention and medical systems while waiting for diagnostic confirmation, the number of intervention hours received without an ASD diagnosis is usually substantially less than many experts recommend.[40-44]

Over the past decade several professional groups, including the American Academy of Neurology,[43] the American Academy of Child and Adolescent Psychiatry,[44] and the American Academy of Pediatrics (AAP)[45,46] have issued guidance on the early detection of ASD. The most recent AAP guidance recommends universal screening of all children at 18 and 24 months of age in addition to developmental surveillance and monitoring. Other agencies have not supported ASD screening: 2011 guidance from the U.K. National Screening Committee does not recommend systematic population screening,[47] nor does 2007 guidance from the Scottish Intercollegiate Guidelines Network (SIGN).[48] These organizations cited a lack of data on PPV in a population setting. However, these recommendations do not reflect all currently available screening research, and the SIGN guidelines are currently under revision but not completed at the time of this report.

After screening positive, some families access diagnostic and treatment services quickly, while other families report significant time (e.g., wait-lists) and financial barriers in accessing evaluation resources.[38,49] Some research reports high levels of parental stress associated with the ASD diagnostic process and advocates additional work to understand the impact of the process on parent functioning.[50] Both delays and demands associated with the ASD diagnostic process may place a burden on the families of children who falsely screen positive, and this is one reason that the screening process in ASD has evolved from the use of single questionnaires with very high false positive rates to two-stage screening approaches that include parent questionnaires and follow up interviews. It is possible that concerns about long waiting lists and/or over-referral of children who meet screening criteria for reasons other than ASD are contributing to low practice of ASD screening in pediatric practice.[51,52,53,54] Increasingly, researchers are attempting to identify mechanisms to train additional community providers (e.g., primary care providers, behavioral providers, educational professionals) to provide timely and accurate diagnosis of ASD and reduce diagnostic wait times. However, variability still exists in families' abilities to access specialized diagnostic assessment resources.

A number of U.S. professional groups and affiliated organizations including the AAP, CDC, and Autism Speaks, have developed materials to help clinicians care for individuals with ASD and related developmental disabilities, including screening and providing follow-up care. This includes information and training on available screeners, coding/billing guidance, and practice support, as well as referrals and resources.

Treatment for ASD

From a health care perspective, the purpose of screening is to identify children for whom effective treatments exist and could make a difference compared to treatments initiated at a later point. Commonly pursued treatments for ASD include behavioral, medical/pharmacotherapy, educational, allied health, and complementary and alternative medicine (CAM) approaches. Our review of the treatment literature is presented below.

Previous USPSTF Recommendation

The USPSTF has not made any prior recommendations on ASD screening.

Chapter 2. Methods

Scope and Purpose

In this review, we systematically reviewed the evidence on screening test accuracy and early treatment effectiveness for ASD in young children, focusing on both benefits and harms of potentially implementing screening for autism spectrum disorder (ASD) in primary care. The research plan for the review was posted for public comment for four weeks and revised based on this input; the final research plan was posted on the U.S. Preventive Services Task Force web site. We explicitly focused on studies that investigated non-selective screening of young (<36 months of age), low risk populations (e.g., universal screening, not specific research paradigms such as infant sibling work). We therefore excluded studies of measures primarily used and studied in populations that have been pre-identified as having ASD, or identified as having some sort of developmental delay either by parents or clinicians, although this is a large body of literature. We included studies of treatments relevant for very young children with ASD (ages 0-4 years) and required that studies include a comparison group and at least 10 total children with ASD. We also reviewed contextual issues from a broader health systems perspective to describe the major issues regarding implementation of screening. **Box 1** outlines screening tools used in included studies. Although additional ASD screeners are published and in use, for this review we only considered tools for which there was available research regarding use in, or use in a close analogue of, unselected primary care screening.

This review will be used to inform U.S. Preventive Services Task Force (USPSTF) recommendations on ASD screening in primary care settings.

Key Questions and Analytic Framework

We determined initial key questions (KQ) in consultation with clinical experts and USPSTF Task Force members. Key questions were also posted to the USPSTF web site for public comment and revised as needed. We addressed the following final key questions:

1. Is screening for ASD in children 12 to 36 months old associated with improved short- and long-term outcomes?
2. What are the performance characteristics (e.g., sensitivity, specificity, positive predictive value, and negative predictive value) of ASD screening tests in children 12 to 36 months old?
    a. Do certain risk factors (e.g., prematurity or having a sibling diagnosed with ASD) modify the performance characteristics of ASD screening tests?
    b. Does the age at which ASD screening is performed modify the performance characteristics of ASD screening tests?
    c. Do other characteristics of the child or family (e.g., intellectual disability, SES, literacy level, insurance status, race/ethnicity, sex, primary language spoken in home, limited English proficiency) modify the performance characteristics of ASD screening tests?
3. What are the harms (e.g., distress, potential misclassification) of ASD screening for the child and family?

4. What is the effect of interventions targeting young children (in preschool and elementary school) on the outcomes of core ASD symptoms, cognitive and intellectual functioning, language and communication skill development, challenging behavior, adaptive behavior, educational placement/achievement, and quality of life for the child and family?
   a. What is the effect of intervention timing (by age and in relation to the establishment of a definitive diagnosis) on treatment outcomes?
   b. What is the effect of severity of ASD (as reported in each study) on treatment outcomes?
5. What are the harms of treatment for ASD in young children?

The analytic framework (**Figure 1**) outlines clinical logic through which children being screened for ASD would proceed in order for there to be positive health outcomes as a result. In short, unselected young children would be accurately screened for ASD and, through appropriate diagnostic evaluation and early treatment, would experience improved developmental and other health-related outcomes, with few or minimal harms as a result of the entire process. Numbers in circles on the diagram illustrate key questions in the process.

Data Sources and Searches

Search Strategy

Databases

A librarian employed search strategies provided in **Appendix A** to retrieve research on screening for ASD in young children and interventions for young children with ASD. All strategies were peer reviewed by a second librarian. Our primary literature search for screening-related studies employed four databases: MEDLINE® via the PubMed interface, PsycINFO (psychology and psychiatry literature), the Educational Resources Information Clearinghouse, and the Cumulative Index of Nursing and Allied Health Literature (CINAHL) database. Our search for intervention studies used the same databases with the exception of CINAHL. In our tests of the strategies, searching CINAHL did not retrieve any unique treatment studies, thus we did not use it for the treatment search.

Our search strategies used a combination of subject heading terms appropriate for each database and key words relevant screening or intervention for ASD. We limited searches to literature published since 2000 to ensure that screening methods and interventions used currently would be represented. We also manually searched the reference lists of included studies and of recent narrative and systematic reviews and meta-analyses addressing ASD screening or intervention in young children with ASD.

Prior Systematic Reviews and Meta-Analyses

We identified systematic reviews and meta-analyses retrieved by the searches for primary literature as well as through scanning the reference lists of included studies. We included summaries of reviews and meta-analyses we rated as good quality (see Quality Assessment below).

Search Terms and Dates

Controlled vocabulary terms served as the foundation of our search for screening-related literature in each database (e.g., MEDLINE vocabulary terms including mass screening, early diagnosis), complemented by additional keyword phrases (e.g., screening, identification). To locate intervention-related studies, we used the search strategy employed in our prior review of therapies for children with ASD.[55] The search used both controlled vocabulary and keyword terms (**Appendix A**). We also limited searches for screening and intervention studies to items published in English and from 2000 to the present. Our searches were done between January and December 2013for screening literature and December 2013 for intervention studies. We updated the MEDLINE search for screening and intervention studies in August 2014. We imported all citations into an electronic database.

Study Selection

Inclusion and Exclusion Criteria

We developed criteria for inclusion and exclusion in consultation with our Medical Officer and USPSTF Task Force members (**Table 1**).

Criteria for Screening Studies

Screening-related studies needed to include at least two individuals (i.e., excluding single case reports) screened for ASD between the ages of 12 and 36 months. We required that studies include undiagnosed populations or populations without suspected developmental delay (i.e., we excluded studies in which the majority of children had an already identified concern about potential developmental delay by parents or clinicians underlying referral for screening/evaluation) or who were already diagnosed with ASD. There is an additional body of research on assessing the ability of ASD screening tests to accurately identify children with ASD in groups of children suspected as having some sort of delay, and as expected, the test performance characteristics are better in these studies than those seen in a general screening population. There is also a body of evidence in which samples are selected that include children with known diagnoses of ASD; these are intended to assess the ability of the screeners to discriminate, but we do not consider that the performance characteristics would apply to the primary care population. None of these scenarios reflects screening in the absence of any concern in the primary care office, in which many children are unrecognized as having symptoms of delay.

We assessed both intermediate and health-related outcomes. Intermediate outcomes included timing of referral and diagnosis and timing of access to intervention. Health-related outcomes included effects on core ASD symptoms, language and communication skill development, and quality of life for the child and caregiver. Screening studies had to take place in primary care or primary care relevant settings.

Criteria for Treatment Studies

To complete the clinical logic of screening to achieve earlier diagnosis and treatment in order to achieve improved health outcomes, we updated a previously published review of treatment for children with ASD,[55] by supplementing it with newer literature meeting the following criteria: 1)Treatment studies needed to include at least 10 individuals with ASD; report on an intervention aimed at young children with ASD (between the ages of 0 and 5 years), and had to include a comparison group. 2) Studies had to evaluate outcomes related to core ASD symptoms, cognitive and intellectual functioning, language and communication skill development, challenging behavior, adaptive behavior, educational placement/achievement, harms of intervention, or quality of life for the child and family. We included studies with any length of followup and in any setting (clinic, home, etc.). We briefly summarize findings of studies addressing interventions under KQ4; detailed results can be found in the full systematic review.

Screening of Studies

Once we identified articles through the electronic database searches, review articles, and bibliographies, we examined abstracts of articles to determine whether studies met our criteria. Two reviewers separately evaluated each abstract for inclusion or exclusion, using an Abstract Review Form (**Appendix B**). If one reviewer concluded that the article could be eligible for the review based on the abstract, we retained it for full text assessment.

Two reviewers independently assessed the full text of each included study using a standardized form (**Appendix B**) that included questions stemming from our inclusion/exclusion criteria. Disagreements between reviewers were resolved by a third-party adjudicator. The group of abstract and full text reviewers included expert clinicians and health services researchers.

Quality Assessment and Data Extraction

Data Extraction and Data Management

The staff members and clinical experts who conducted this review jointly developed the evidence table, which was used to summarize data from the studies. We modeled the table on USPSTF methods guidelines and designed the table to include issues of study design, descriptions of the study populations, description of the screening process or intervention, and baseline and outcome data on constructs of interest.

One team member initially entered information into the evidence table. Another member of the team also independently reviewed the articles and edited all initial table entries for accuracy, completeness, and consistency. The full research team met during the article extraction period and discussed issues related to data extraction (e.g., optimal level of detail in the description of the screening technique or intervention, determining key population characteristic to include). In addition to outcomes related to screening performance and intervention effectiveness, we extracted all data available on harms. Harms encompass the full range of specific negative effects, including the narrower definition of adverse events. The final evidence table is presented

in **Appendix C**. Studies are presented in the evidence table alphabetically by the last name of the first author within each year.

Quality (Risk of Bias) Assessment of Individual Studies

Screening Studies

We assessed the quality of screening studies using design-specific quality criteria based on the USPSTF methods.

Intervention Studies

We assessed the quality of intervention studies using methods previously developed for systematic reviews of interventions for children with ASD.[55] We evaluated the quality of studies in the domains below using specific questions to evaluate a study's conduct. We rated each domain individually and combined them for an overall quality level. Three levels were possible: good, fair, and poor.

Study Design

1. Did the study employ a group design (have a comparison group)?
2. Were the groups randomly assigned?
3. If no, was there an appropriate comparison group?
4. If yes, was randomization done correctly?

Diagnostic Approach

1. Was a valid diagnostic approach for ASD used within the study, or were referred participants diagnosed using a valid approach?
    a. A clinical diagnosis based on the DSM, in addition to the Autism Diagnostic Interview-Revised (ADI-R) and/or Autism Diagnostic Observation Schedule (ADOS) assessments.
    b. A combination of a DSM clinical diagnosis with one other assessment tool; or the ADOS assessment in combination with one other assessment tool.
    c. Either a clinical DSM-based diagnosis alone or the ADOS assessment alone.
    d. Neither a clinical DSM-based diagnosis nor the ADOS assessment

Participant Ascertainment

1. Was the sample clearly characterized (e.g., information provided to characterize participants in terms of impairments associated with their ASD, such as cognitive or developmental level)?
2. Were inclusion and exclusion criteria clearly stated?
3. Do the authors report attrition?
4. Were characteristics of the drop-out group evaluated for differences with the participant group as a whole?

Intervention Characteristics

1. Was the intervention fully described?
2. Was treatment fidelity monitored in a systematic way? (for non-medical interventions)
3. Did the authors measure and report adherence to the intended treatment process? (for medical interventions)
4. Did the authors report differences in or hold steady all concomitant interventions?

Outcomes Measurement

1. Did outcome measures demonstrate adequate reliability and validity (including inter-observer reliability for behavior observation coding)?
2. Were outcomes coded and assessed by individuals blinded to the intervention status of the participants?

Statistical Analysis

1. For RCTs, was there an intent-to-treat analysis?
2. For negative studies, was a power calculation provided?
3. For observational studies, were potential confounders and effect measure modifiers captured?
4. For observational studies, were potential confounders and effect measure modifiers handled appropriately?

Finally, we assessed the quality of systematic reviews and meta-analyses using the AMSTAR tool.[56] For all types of studies, two reviewers independently assessed quality, with final decisions made via discussion to reach consensus or by third party adjudication by a senior methodologist as needed. We report individual quality assessments for each study in **Appendix D**.

Determining Quality Levels

Screening Studies

We determined quality ratings for screening studies based on USPSTF methods and criteria for ratings for diagnostic accuracy studies.[57] Our criteria were as follows:

Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 1,000) patients with and without disease, includes participants drawn from the general population and follows at least a random sample of screen negative participants.

Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size.

Poor: Has fatal flaw such as: Uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample

size.

An important consideration in assessing quality of the screening accuracy studies was the methodology employed for handling screen-negative participants. Given the prevalence rate of 1 in 68, as well as the resource-intensive nature of conducting the gold standard evaluations (i.e., in-person, time-intensive diagnostic assessments by skilled behavioral professionals sometimes lasting a full day or more) in screen negative children, almost no study has followed a large enough sample to truly assess false negatives on a population level. Given the pervasive and ongoing nature of developmental disorders, however, false negatives are primarily a concern among borderline cases, such as those that fail one portion of the screening process. Rather than follow up with all screen negatives, some studies followed up with borderline cases that failed a portion of the screening process (e.g. failed the screener but not the interview). As such, we allowed varying methodologies for attempting to assess false negatives among those who partially failed screening to qualify in determination of a "good" study if they met other criteria as a good study.

Intervention Studies

We assessed each domain described above individually and considered the individual ratings to determine an overall quality assessment of good, fair, or poor. We required that studies receive positive scores questions related to study design and diagnostic approach to be considered good quality. Scores were calculated first by domain and then summed and weighted as described in **Table 2** to determine overall study quality. Studies could receive up to two points on the domains of study design, diagnostic approach, participant ascertainment, and intervention, and up to one point on the domains of outcome measurement and statistical analysis.

Data Synthesis and Analysis

We determined that a meta-analysis of screening studies would be inappropriate and unnecessary. We assessed attrition at each step in the process and assumed non-differential loss to followup to impute diagnostic yield and other outcomes if there had been complete followup for key studies. We analyzed results of studies qualitatively, summarizing them in tables and in text. We summarized only those systematic reviews rated as good quality.

USPSTF Involvement

We worked with USPSTF liaisons at key points in the review process to develop the analytic framework and key questions, to address methodological decisions on applicable evidence, and to resolve issues of scope for the review. The review was funded by AHRQ under a contract to support the work of the USPSTF. AHRQ staff provided oversight for the project, reviewed the draft report, and assisted with external review of the draft report.

Chapter 3. Results

Literature Search

Screening Studies

We identified 3,469 citations potentially addressing screening for autism spectrum disorders (ASD). We excluded 3,050 publications at the abstract review stage and 435 at the full text stage (**Appendix E**). We summarize results from 17 unique studies (reported in 22 publications; one publication reports two separate studies) meeting our inclusion criteria in this review. **Figure 2** outlines the disposition of screening studies in primary care settings. We also summarize information about studies identified for the treatment-related key questions (Key Questions [KQ] 4-5).

Among the 17 unique primary care screening studies described in this review,[39,58-78] we rated five as good quality,[39,61,67-71,75,76] 10 as fair quality,[58,59,62,63,65,66,72-74,78,79] one as fair to poor quality,[77] and one as poor quality.[64] Most studies (n=6) were conducted in Europe,[62,64-66,75,78] six in the United States,[39,61,63,67-71,76,79] three in Japan,[59,72,77] and one each in Australia and Israel.[58,73,74] Participant ages ranged from 4 to 36 months, and studies screened between 583 and nearly 35,000 children. **Box 2** summarizes the characteristics of these included studies, and **Appendix C** includes evidence tables for each study.

Among the excluded studies were a number that assessed screening in children who had already been identified as having some sort of concern for developmental delay and therefore did not reflect an unselected primary care screening population.[80-100] Most studies included children with known developmental delay or those referred to specialized centers for suspected developmental issues; four papers, reporting on data from populations that likely overlap though the reporting is not clear, assessed children already receiving early intervention services for an unspecified developmental delay.[84-87] Two studies evaluated screening in younger siblings of children with ASD but were not in a primary care setting or population.[90,94]

Intervention Studies

We identified 2,639 citations and abstracts (**Figure 3**). We excluded 2,012 studies at abstract review and assessed the full text of 627 studies. Among these, 55 publications, comprising 42 unique studies, met our criteria for intervention studies and were rated as good or fair quality. These studies included 26 RCTs (nine good and 17 fair quality), five non-randomized trials (one good and four fair quality), 10 prospective studies (two good and eight fair quality), and one retrospective cohort study (fair quality). Studies were conducted in the United States (n=21), the United Kingdom (n=6), the Netherlands (n=3), Norway and/or Sweden (n=4), Australia (n=2), Canada (n=2), Israel (n=2), Belgium (n=1), and Italy (n=1). Studies used early intensive behavioral and developmental interventions, all based to varying degrees on Applied Behavioral Analysis (ABA). We grouped them into three categories: direct provision to the child (n=17); incorporating parent training (n=13) or play/interaction-focused (n=12).

Key Question 1. Is Screening for ASD in Children 12 to 36 Months Old Associated With Improved Short- and Long-Term Outcomes?

We did not identify studies that directly compared screening (vs. no screening) in terms of longer-term health and social outcomes.

Key Question 2. What Are the Performance Characteristics of ASD Screening Tests in Children 12 to 36 Months Old?

Checklist for Autism in Toddlers (CHAT)

Key Summary Points

- In one good quality study, the sensitivity of the CHAT for first-round screening conducted by nurses was 35.1 percent (specificity 99.9%). For the second screen of children who screened positive on the first round, the positive predictive value (PPV) at a high risk threshold was 83.3 percent and 58.8 percent using a medium risk threshold. Overall, the tool was specific (>95%), but sensitivity for detecting ASD was low (21.3%).
- Use of the CHAT resulted in substantial under-identification of ASD (i.e., low sensitivity and modest PPV) in a low risk population. Due to this concern it has been in limited use in the U.S. since publication of these findings and is largely replaced by the M-CHAT.

Description of Measure

The Checklist for Autism in Toddlers (CHAT) was one of the first formal screeners developed in the UK for early identification of ASD and was designed in an attempt to identify autism at 18 months. The CHAT assesses pretend play, pointing, and gaze monitoring by both parent report and practitioner observation via direct testing.

Detailed Analysis

One good quality study (Baird 2000) followed up on a 12-month birth cohort of children in the UK who were screened using the CHAT at 18 months of age.[75] Additional screening took place at 3 and 5 years of age using the Checklist for Referral and the Pervasive Developmental Disorders Questionnaire. The population was then reexamined at 7 years of age to determine the performance characteristics of the CHAT. This was accomplished through a broad examination of available data potentially identifying ASD in this population, including examining referrals to clinical centers and medical/educational/social service and other available records. Children identified at risk by any of these screenings were directly assessed by a research team and assigned ICD-10 diagnoses.

A sample of 16,235 toddlers was administered (13,694) or mailed (2,541) the CHAT at 18 months of age by their primary health care providers, who were primarily nurses (1-stage). Those who screened positive were re-screened one month later by the research team (2-stage), with children with severe developmental delays excluded at the discretion of the home visitor. At both stages toddlers were placed into a high risk (failure of five critical items) or medium risk category (failure of item A7 and B4: both items regarding pointing). During the first round of screening, 38 children met the threshold for high risk and 369 met the threshold for medium risk. The sensitivity of the CHAT for this first round of screening using both medium and high risk thresholds was 35.1 percent (specificity 99.9%). During the second round of screening, 12 children continued to meet the threshold for high risk and 22 continued to meet for medium risk. Due to resource constraints, investigators rescreened only half of the medium risk cases. For the second screen, the positive predictive value (PPV) of the high risk threshold was 83.3 percent and 58.8 percent for the medium risk threshold. Overall, while the measure was quite specific (>95%) the sensitivity for detecting ASD was low (21.3%). The investigators hypothesized that false negatives may have resulted from the following concerns: 1) asking parents if they had "ever" demonstrated a behavior vs. "rarely" doing so might be a less sensitive measurement strategy, 2) identification of communicative pointing by parents at young ages may be challenging, and 3) the potential of a late-onset and/or regressive form of ASD manifesting more clearly at later ages. Important limits of the study included only 40 percent representation of population cohort, exclusion of children with "handicaps," screening conducted via researchers vs. clinicians, and lack of reliability data on the instrument.

An additional poor quality study (VanDenHeuvel 2007) examined the use of the CHAT at 18 month well check visits with public health nurses.[64] The sample included 2,117 18-month olds (1,029 females, 1,088 males). Due to methodological limitations, true sensitivity, specificity, negative predictive value (NPV) and PPV were not able to be calculated. However, the identification of only seven children from this cohort of is in-line with concerns about the sensitivity of the CHAT as a standalone instrument documented in previous work.[75]

Modified Checklist for Autism in Toddlers (M-CHAT), M-CHAT With Follow-Up Interview (M-CHAT/F)

Key Summary Points

- Two good and four fair quality studies assessed the use of Modified Checklist for Autism in Toddlers (M-CHAT) and the use of the M-CHAT Follow-Up Interview (M-CHAT/F) in children 12 to 36 months of age without previously identified symptoms of developmental delay.
- Positive predictive value using the M-CHAT questionnaire and follow up interview as recommended was 54 percent in a large, good quality study of over 18,000 children in the United States.
- In that study, nearly all 171 screen positive children who underwent a diagnostic evaluation received final diagnosis of either ASD (n=92) or some other developmental concern (n=75). Four received a final diagnosis of normal development.
- Not using the follow up interview component results in substantial over identification, with approximately 10 percent initially identified as screen positives.

- Attrition rates were approximately 24 percent between the initial screen and follow up and 39 percent between follow up interview and diagnostic evaluation.
- There are not currently good estimates of how many children with ASD screen negative on the M-CHAT, although validation procedures and population estimates suggest this number may be substantial.

Description of Measure

The M-CHAT is a 23-item parent-report (Yes/No) ASD screening tool initially designed for children 16 to 30 months of age. It differs from the original CHAT in that: 1) it covers a more extensive set of developmental domains (e.g., sensory, motor) and 2) it relies exclusively on parent report without clinical observation. To address concerns about potential over-identification of risk based on parent report alone, the authors of the instrument later formally described and studied a specific follow-up interview (M-CHAT/F). This followup interview is conducted in-person or via phone. Current use of the instrument is the combination of the parent report with the follow-up interview (for children in the moderate risk range), with the authors explicitly noting that questionnaire use alone probably results in increased false positive identification of ASD risk. This tool is the most common approach to ASD screening in the United States and is recommended by the American Academy of Pediatrics (AAP).

Overview

We identified six studies—two good (one reported in multiple publications) and four fair quality studies (reported in three publications)—assessing the use of Modified Checklist for Autism in Toddlers (M-CHAT) and the use of the M-CHAT Follow-Up Interview (M-CHAT/F) in children 12 to 36 months of age without previously identified symptoms of developmental delay[59,62,67,69-71,77,78,101] One good quality study is reported in five publications and includes data on the initial sample of children for whom reliability and validity of the measure were assessed. These publications provide results on the use of the M-CHAT alone and use of the M-CHAT and M-CHAT-F. Several of the studies include information drawn from both primary care (e.g., low risk) and clinical referral (e.g., high risk) samples. Another publication includes two separate studies.[78] One study was conducted in the United States,[67,69-71,101] two were conducted in Japan,[59,77] two (reported in one paper) were conducted in Spain,[78] and one was conducted in Sweden.[62]

Detailed Analysis

One good quality study (Chlebowski 2013) reported in multiple publications[67,69-71,101] included 18,989 toddlers 18 to 24 months of age (Mean=20.4, SD=3.1; 9,388 female, 9,601 male; 6,184 White/ non-Hispanic, 1,186 Nonwhite) who were screened at pediatric well-child visits in two regions of the United States (Georgia State University and University of Connecticut catchments). Results for some of the toddlers had been previously reported in other studies and were compiled with additional participants here. Toddlers were excluded if they had already received an ASD diagnosis or risk classification, had a severe sensory or motor disability (e.g., blindness, deafness) or if their parents were not fluent in English or Spanish.

M-CHAT forms were distributed at participating pediatric offices and sent to research staff for scoring. Parents of toddlers who screened positive on the M-CHAT (screening positive on 2 of the 6 critical items or any 3 of the 23 item measure) were called and completed the M-CHAT/F over the phone. Children who screened positive on the M-CHAT/F were offered a free comprehensive diagnostic evaluation. This evaluation, conducted by a licensed clinical psychologist or developmental pediatrician plus one or more doctoral trainees/research staff, included the Autism Diagnostic Observation Schedule (ADOS) and, for some participants, the Autism Diagnostic Interview-Revised (ADI-R), as well as cognitive and adaptive measures. Diagnoses were made based on the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*.

Of the 18,989 children screened with the M-CHAT, 1,737 (9.1%) screened positive. About three quarters (1,295; 74.6%) of these children went on to complete the M-CHAT/F. After this screen (M-CHAT/F), 272/1,295 children (1.4% of total screened sample) continued to screen positive and were offered evaluations. Evaluations were completed on 165 of the children who screened positive on the M-CHAT (questionnaire) and M-CHAT/F (follow-up interview) as well as 6 children who screened positive on such a high number of items on the M-CHAT (≥7) that the M-CHAT/F was not conducted (total n evaluations=171). Evaluations were also completed on 36 children who were identified as potential missed cases. These children were either "red flagged" by their pediatricians due to concerns or screened positive on a separate screening measure. Although all pediatricians were instructed to flag ASD concerns, these screeners and flags were not used systematically for the entire sample.

Out of the 171 evaluations completed on children considered as screening at risk, 92/171 received a diagnosis of ASD, 60/171 received a non-ASD DSM-IV diagnosis (21.6%, n=37 developmental delay, 10.5%, n=18 language disorder, 2.9%, n=5 other behavioral diagnoses), 15/171 (8.8%) were noted to have developmental concerns but did not receive a diagnosis, and four of 171 (2.3%) were typically developing. The positive predictive value (PPV) for identifying ASD with the M-CHAT and M-CHAT/F was 54 percent. In the previous publications that included subsamples of the overall group of toddlers,[67,69-71] PPV for identifying ASD including a follow-up procedure (M-CHAT/F) ranged from 57 to 79 percent. The PPV for identifying any developmental concern (ASD, other diagnosis or developmental concern) in this population is 98 percent.

Of the 36 children receiving evaluations due to pediatrician flags as potential missed M-CHAT screen positives, six received an ASD diagnosis, 18/36 received non-ASD DSM-IV diagnoses (50%), nine were identified with developmental concerns (25%), and three were typically developing (8.3%).

The methods used for following screen negatives do not permit calculations of sensitivity, specificity, or NPV. The total number of children successfully identified by this screening study (0.48%) at an average of 20 months is significantly below the currently estimated population prevalence (1.47%), but expected prevalence in the current sample would likely be lower because it excluded children identified early by parent or clinical concern, or lost to follow up (40% of screen positive cases did not receive diagnostic evaluation). The investigators evaluated different scoring systems post-hoc and found that a revised scoring algorithm (≥3 total and no

use of critical scores) would yield the best identification of screen-positive children. Additionally, 82.2 percent children with a score of ≥7 continued to screen positive when administered the M-CHAT/F, which may suggest that the follow-up interview could be unnecessary for very high M-CHAT questionnaire scores.

An additional fair quality study conducted in the United States (Miller 2011), screened toddlers ages 14 to 30 months at all scheduled medical appointments (well-child, sick, follow up, injection visit) during a 6-month period at a large community based pediatric practice.[63] Caregivers were given an M-CHAT, the Infant Toddler Checklist (ITC), or both to fill out when they checked in. The ITC is a 24-item developmental screener of language and communication that also includes a yes/no question about whether the caregiver is concerned about their child's development. Children failing the M-CHAT (≥3 total or ≥2 critical items) or the ITC (below 10th percentile on Social, Symbolic, or Total scores) were then referred to research staff. If the child failed the M-CHAT, the M-CHAT follow-up interview was completed. If the child failed the ITC, items were reviewed over the phone by trained graduate students. Children administered either test who had a confirmed positive screen based on interview were invited to a full evaluation using the ADOS, the Mullen Scales of Early Learning, and the Vineland Adaptive Behavior Scales. During the study, 990 toddlers had appointments at the clinic, and 796 were screened. Of these, 192 (24%) screened positive on either the M-CHAT or the ITC, with 47 of 192 (24.5%) continuing to screen positive after the follow up call/interview. Ninety-eight of 192 screened negative (51%), and 47 of 192 (24%) either declined or were not reached by telephone. Of those that were not lost to follow up or that declined (n=145), the proportion screening positive (n=47) was 32.4 percent.

The 47 children screening positive, plus two children who screened negative but whose parents raised developmental concerns after the children were screened (total n=49), were referred on for in-person evaluations. Thirty of these 49 received evaluations, as 12 did not show up and seven declined. After full evaluation, 13 of 30 children were diagnosed with or at high risk for ASD (43%), 16 had delays other than ASD (53%), one was typically developing (3%). However, of the 13 children who were identified as having ASD or at high risk, three had been previously diagnosed with ASD and should thus be excluded for calculations of the performance characteristics of the screener (therefore, the total N for calculating performance characteristics is 27, of which 25 screened positive). Of the 10 new diagnoses, three screened negative on the questionnaires and were therefore identified using means other than the screening process. In one of those cases, the child subsequently screened positive on the M-CHAT and is therefore considered to be a screen positive-case positive; thus, there were a total of 8 children who both screened positive and were diagnosed as positive. PPV for the combined use of ITC plus interview or M-CHAT plus interview as an ASD screener was, therefore, 32 percent (8/25). No information on the PPV of either tool alone was presented in this study.

As a result of screening, 60 percent of identified cases of ASD were found before parents expressed concern, 50 percent were identified before pediatrician concern, and in only 20 percent of cases were both the parent and pediatrician concerned before the screener identified a problem. These percentages are not mutually exclusive. Two children who did not screen positive in this process were flagged and referred on for further evaluation by providers based on clinical concerns. Both of these cases showed early signs of ASD during the in-person evaluation.[63]

Four additional publications reported on the use of the M-CHAT/F in different language groups in different countries (Japan, Sweden, Spain).[59,62,77,78] All studies translated and back-translated the screener in accordance with guidelines set by authors, and made cultural-linguistic alterations to certain items to preserve item intent. Such subtle alterations and applications to varied cultural groups may affect reported psychometric properties of instrument use. However, to be comprehensive, and because even within the U.S. universal screening requires linguistic, if not cultural adaptation, we report results of these studies here.

One fair to poor quality, population-based study (Inada 2011) examined the use of the M-CHAT-J questionnaire without the follow up interview in a group of 1,187 (575 female, 612 male) children in whom data were available at the 18 month and 36 month checkup in Munakata City, Japan.[77] Twenty of the 1,187 children (2%) were diagnosed with an ASD. These diagnoses were made based on developmental history, clinical assessment and *DSM-IV* criteria. Diagnoses were confirmed by consensus between two psychiatrists and a psychologist with experience in child development. Several scoring cutoffs were explored, as well as the use of a shorter 9-item version of the M-CHAT-J. A cutoff of ≥2 on the 23-item version resulted in a sensitivity of 75 percent, specificity of 89.3 percent, PPV of 10.7 percent and a NPV of 99.5 percent. The traditional ≥3 item scoring resulting in sensitivity of 55 percent, a specificity of 96 percent, a PPV of 19 percent, and NPV of 99 percent. The low PPV across groups reflects a large number of false positives in this sample where no embedded follow-up procedure or interview was clearly specified.

Another fair quality population-based study (Kamio 2014) examined the use of the M-CHAT/F-Japanese in 2,113 toddlers attending their 18-month well-check in Fukuoka City, Japan.[59] Parents of 1,851 children completed the screening and had follow-up data available. It is unclear from the description whether any toddlers were lost to followup and whether the number described represents only those with followup, or whether the entire population did receive followup assessment. Modified M-CHAT/F-J scoring cutoffs were used that included labeling additional items as "critical" and lowering the threshold for ASD concern on the parent checklist. Original scoring criteria were used for the followup interview. Screen positive children were invited for diagnostic evaluations by a team of child psychologists, psychiatrists, and nurses at age 2. These children were then invited back for full evaluations at ages 3, 4, and 5. The investigators created a 20-item checklist in order to follow-up potential false negatives via well visits at age three (i.e., children failing this questionnaire were also offered evaluation along with screen positive children). Among the 1,851 children screened, 319 (17%) screened positive using the M-CHAT and 1,532 (83%) screened negative. One hundred and ninety-five children among the 319 screening positive completed the follow-up interview, and 124 did not. Among the 195 of 319 who completed the follow-up interview (61%), 44/195 (23%) continued to screen positive, and 151/195 (77%) screened negative. Twenty children (45%) among the 44 screening positive on the follow-up interview were diagnosed with ASD.

In addition, because most children attended day care or local kindergarten, the investigators sought additional information available at those locations to identify cases in the 3- to 5-year range. In total, 51 children were identified with ASD through the combination of screening and case-seeking. Forty-two of these children had participated in research screening, with 20 screening positive via the modified M-CHAT/F-J scoring procedure.

One publication (Canal-Bedia 2011) reported two separate fair quality studies (one considered a validation study and one a reliability study) completed in Madrid, Spain and using the M-CHAT/F translated into Spanish.[78] The first study (validation of the translated M-CHAT) included 2,480 children ages 18-36 months (1,163 female, 1,254 male) who were recruited during their mandatory 18-month vaccination appointment or their 24-month well-child visit. This population also included 63 children who were considered high risk for ASD and who had developmental ages of 18 to 24 months (maximum chronological age=48 months, range not provided). Primary care pediatricians and nurses distributed the M-CHAT, which was completed by parents, and psychologists with training in child development conducted the followup interview with parents of those children who screened positive. Out of the 2,480 children screened, 429 screened positive on the initial questionnaire (17%), and 86/429 (20%) continued to screen positive following the follow up. All 86 children who screened positive underwent a diagnostic assessment based on DSM-IV criteria. As a result of this assessment, 23/86 children (27%) were identified as having ASD. Nineteen of these cases were from the high risk sample. The PPV of the M-CHAT/F in this study was 27 percent.

The second study (reliability study) reported in the same publication[78] was a population based study in Madrid and included 2,055 children (949 female, 1,106 male) 18 to 36 months old attending their vaccination or well-child visit. The procedures were the same (i.e., primary care pediatrician or nurse distributing the questionnaire, psychologist completing the follow up interview). Out of the 2,055 screened, 336 (16%) screened positive on the M-CHAT, and 31/336 (9%) continued to screen positive following the M-CHAT/F. All 31 completed a diagnostic assessment, and 6/31 (19%) were ultimately diagnosed with an ASD, resulting in a PPV of 19 percent. This was a population with no known cases at high risk initially.

Another fair quality study (Nygren 2012) screened 3,999 children (1,912 female, 2,087 male) at their 2.5 year checkup at child health centers in Gothenburg, Sweden.[62] This study is part of a larger effort implement a screening program in Gothenburg. The authors estimated that 80 percent of all 2.5 year olds (approximately 5,000) in the area were screened. This analysis focused on those children who were younger than 24 months or older than 36 months. The M-CHAT questionnaire and followup interview were translated into Swedish with the guidance of the original authors to ensure integrity of the instrument. The questionnaire was mailed to the family before their appointment. Nurses (who had been trained on scoring the M-CHAT/F) administered the followup interview in person to those children who had screened positive on the initial questionnaire. These nurses were also trained on developmental milestones, early signs of ASD as well as a Joint Attention Observation of Toddlers (JA-OBS) screen. The JA-OBS consists of observations by the examiner of child response to name, eye contact, response to examiner's direction of attention (finger pointing), child's use of finger pointing to direct attention, and pretend play. Failure of ≥2 items constitutes a positive screen. Both screening approaches were used for all children.

If a child screened positive on either the M-CHAT/F or JA-OBS or if there was a suspicion of ASD for other (undefined) reasons, a separate appointment was made with a pediatrician to take family medical history and screen for other possible medical problems. The pediatrician then referred children who still had a high level of ASD concerns to a specialist. Of 3,999 children screened using the M-CHAT/F and J-OBS, 64 were considered at risk on either tool: 62 detected

via the screening process and two referred from speech pathologists. Ten children who screened positive on any mechanism were lost to follow up.

Of the total of 64 children screening positive, 36 were identified using the M-CHAT/F, 33 of whom received a diagnosis of an ASD after neuropsychiatric confirmatory exam. The PPV for the M-CHAT/F was therefore 91.7 percent. Forty children screened positive on the JA-OBS; 37 of these children ultimately received a diagnosis of an ASD (92.5%). The PPV for the JA-OBS was 92.5 percent (95% CI: 79.6 to 98.4%). Of the 54 who had a positive screen through any mechanism and went on to a diagnostic assessment, 51 had been screened with both the M-CHAT/M-CHAT/F and the JA-OBS, and 45 of the 48 who ultimately received an ASD diagnosis had been screened with both tools. Of these, 43 screened positive on either one or both of the tools, yielding an overall PPV for screening of 89.6 percent (95% CI: 77.3 to 96.5).

Modified Checklist for Autism in Toddlers, Revised With Follow-Up (M-CHAT-R/F)

Key Summary Points

- The M-CHAT-R/F is a two-part screening process assessed in one good quality U.S. study including 16,115 toddlers between 16 and 31 months of age.
- Using the M-CHAT-R/F identified significantly more children with ASD than the M-CHAT/F (67 per 10,000 vs. 45 per 10,000, p=0.003).
- Use of the M-CHAT-R/F procedure (e.g., questionnaire and formal interview) within diverse community based primary care settings had a positive predictive value of 48 percent.
- Eighteen percent of participants were lost to followup between the initial screening and followup interview, and 57 percent were lost between the followup interview and diagnostic evaluation.
- Diagnoses included ASD (n=123), global developmental delay (n=61), language delay (n=25), other unspecified diagnosis (n=1), no diagnosis (n=30), typically developing (n=23).
- It is unclear as to how many children with ASD screen negative on the M-CHAT-R/F at young ages or how the procedure operates in high-risk samples.

Description of Measure

The M-CHAT-R/F is a two-stage screening process. Parents complete a 20-item parent-report (Yes / No) ASD screening tool. Modifications from the original M-CHAT questionnaire include 1) deletion of items with reported poor performance, 2) reordering of items to guard against agreement bias in reporting, 3) simplification of language to improve comprehension, 4) and provision of specific examples to anchor behavioral reporting. If children screen positive, parents then participate in a structured follow-up interview to obtain additional examples to assess risk. This is designed to be able to be completed by a pediatric extender or support staff. Scoring of the M-CHAT-R/F is based on total scores from parent report and/or interview (i.e., scoring not based on critical item scores). This modified tool became available for free use in January 2014.

Detailed Analysis

The most recent study in the United States of population level screening for autism used the M-CHAT R/F, a revised version of the M-CHAT with follow up (Robins 2014).[61] The study included 16,115 toddlers 16 to 31 months of age (Mean=20.95 ± 3.3; 7,570 female, 7,793 male; 6,184 White/ non-Hispanic, 1,186 Nonwhite) who were screened at 85 pediatric clinics (41 metropolitan Atlanta, 44 University of Connecticut catchment areas).[61] Toddlers were excluded if they had already received an ASD diagnosis or risk classification, had a medical condition that precluded evaluation, or if their parents were not proficient in English.

M-CHAT-R forms were distributed during 18 or 24-month well care visits. Pediatricians also were asked to indicate any clinical concern about ASD on top of the screening form. Research staff scored the completed forms and completed followup telephone interviews. Toddlers who failed the screening protocol (questionnaire plus followup) or whose physician had indicated concerns were offered a free diagnostic evaluation. This evaluation, supervised by a licensed clinical psychologist or developmental pediatrician, included the ADOS, the Childhood Autism Rating Scale-2, the Toddler Autism Symptom Interview, and cognitive and adaptive measures. Diagnoses were based on the *DSM-IV* using all available information.

Of the 16,115 children screened with the questionnaire, 1,155 (7.2%) initially screened positive. Of these 1,155, 946 (81.9%) went on to complete the followup interview, thus completing the screening process. After this screen, 348/946 children (2.2% of total screened sample, and 30% of those who had initially screened positive) continued to screen positive and were offered evaluations. Two hundred twenty-one of the 348 children (64%) who screened positive on the combined questionnaire and follow up interview were evaluated for ASD. About half received an ASD diagnosis (n=105/221, 47.5%). In addition, 79/221 (35.7%) had "other delays," 25/221 (11.3%) had developmental delays not associated with a diagnosis, and 12/221 (4.5%) were typically developing. The authors calculated a PPV for any diagnosis as .946 (95% CI: 0.92 – 0.98) but do not indicate what the specific "other delays" diagnosed were.

A small, quasi-random sample of children who had screened negative at the Atlanta site (n=375) were given a second screener, the Screening Tool for Autism in Two-Year Olds (STAT). The STAT is a brief (20 minute) interactive ASD specific screener administered by a trained clinical researcher designed to yield an overall ASD risk classification based on assessment of core communication and play skills. Children who had initially scored positive on the M-CHAT-R but negative on the followup were most heavily recruited for this assessment. Of the 375 children recruited, 20 screened positive on the STAT and were sent for further evaluation. Six (30%) of these 20 children went on to receive ASD diagnoses. With regard to physician ASD concerns (i.e., box for primary care provider to indicate ASD concerns on M-CHAT), 45 of 64 identified cases of concern attended an evaluation with 42 diagnosed with ASD (n=30) or other concerns (n=12).

The criteria for assessing screen failure based on the followup interview were altered during the course of the study. Initially a total cutoff score of ≥3 on both the questionnaire and followup interview was used. However, five of seven cases that had screened negative on this basis, and who had a cutpoint of two but were identified as a concern by clinicians, were found to have

ASD, so the cutpoint was lowered to 2. Performance characteristics after the change showed improved sensitivity. The authors also noted that all children scoring above 7 on the parent questionnaire ultimately were identified with ASD or an actionable developmental concern, suggesting that referral for diagnostic evaluation after one-step screening with M-CHAT-R may be appropriate for such cases.

The total number of children successfully identified by this screening study (rate of 0.65%) at an average of 21 months is significantly below the known population prevalence (1.47%), but expected prevalence in the current sample would likely be lower because children identified early by parent or clinical concern, or lost to follow up (29% of screen positive cases were not followed up) were excluded. If one assumed no differential attrition, projected prevalence of screen detected ASD would be 92 percent at 20 months.

Infant Toddler Checklist (ITC)

Key Summary Points

- The ITC is a general developmental screener for communication development, not an ASD specific tool. Studies have significant limitations.
- Two good quality studies conducted in the United States and including a total of 15,864 children between the ages of 6 and 24 months assessed the ITC.
- Use of the ITC questionnaire in isolation may designate ≥10 percent of the total low-risk population as screening positive for ASD, many of whom clearly will not go on to an ASD or other developmental diagnosis.
- When considering the ITC's original purpose to identify all types of delays (including ASD, language delay, developmental delay, and Other), PPV was 75 percent.
- Assessing only the ASD diagnoses results in a PPV of 18 percent.

Description of Measure

The ITC is a component of the Communication and Symbolic Behavior Scales Developmental Profile (CSBS).[102] The CSBS is an assessment methodology for attempting to identify and gauge a broad range of developmental risks in young children. In contrast to the ASD-specific measures of the CHAT and M-CHAT, the ITC is designed as a screener for general developmental concerns to identify communication delays and disorders, including ASD. This parent report questionnaire includes 24 items focused on social communication milestones rated on a Likert scale, and an open-ended question about current concerns. The ITC has been standardized in a normative sample; its results can be reported as screening cutoffs (e.g., risk cut-offs) or as standard scores for infants and toddlers between 6 to 24 months.

Detailed Analysis

One good quality study examined the effectiveness of the ITC as an ASD screener for young children in primary care settings (Pierce 2011).[39] Children (n=10,479) across San Diego County, California were screened at their 12-month well-child pediatric visit (mean age at screening= 12.54 months, range=10.08-15.97) using the ITC. Children who failed the screener, and were

further evaluated using medical chart review along with a selected sample of children who passed the screener (screen negative), were referred for a complete developmental evaluation. Participants were recruited via a flyer with study personnel not allowed to contact those failing the screening directly. The diagnostic evaluation included the ADOS and the Mullen Scales of Early Learning. Children were re-evaluated every 6 months until the age of 3. Children eventually diagnosed with ASD had a mean 6.0±2.5 visits. Re-evaluations included a re-administration of the ITC, ADOS, and the Mullen Scales.

After initial one-year administration, 1,318/10,479 children failed the screener (12.6%, with failure defined as scoring below the 10th percentile) and 346/1318 (26.3%) were referred for further evaluation after chart review. The authors do not explain why the majority (73.7%, n=972) of children failing the screening were not referred for evaluation. Of the 346 referred, 208 enrolled in the study, but 24 dropped out or did not respond to follow-up requests, leaving 184 screen positive children who completed the initial screener as well as all of the follow-up appointments. The study also included 41 children who initially passed the screener at the 12-month well child check visit and were considered controls (total n=225). Five children were initially identified as having an ASD, but this diagnosis was later removed. The authors did not include these children as false positive results in their own calculations of performance characteristics. Of the 184 screen-positive children who received diagnostic evaluation, 32 (17%) were diagnosed with an ASD, 56 (30%) had a language delay, nine (5%) had a developmental delay, 36 (20%) had another diagnosis (Other), 5 had prior ASD diagnoses as noted above and 46 (25%) were false-positive results (no additional detail provided). All 41 children who passed the screener were identified as typically developing. When considering the ITC's original purpose to identify all types of delays (including ASD, language delay, developmental delay, and Other), PPV was calculated to be 75 percent. When considering the ITC as an ASD-specific screener, PPV was calculated to be 17.4 percent.

The authors also tracked treatment engagement subsequent to screening, with 100 percent of toddlers with ASD and developmental delay and 89 percent of the sample with language delay referred for early intervention as soon as delay was documented via ITC scores and clinical concerns. Treatment on average began at approximately 17 months of age, with children in the ASD group receiving 11.5 hours per week of treatment as compared with those in the language delay group (1.9 hours per week).[39]

In a second good quality study (Wetherby 2008)[76] a cumulative community sample of 5,385 children between the ages of 6 and 24 months was recruited from a variety of health and child care settings in the United States. Service providers were encouraged to administer the ITC each time the child was seen, so some children had multiple assessments. If children screened positive on the ITC or if their parents raised concerns, they were assessed using the Behavioral Sample of the CSBS-DP. Screening data were then linked to ongoing ASD prevalence studies being conducted by this research group to identify children that received ASD diagnoses using 1) scoring from the Behavioral Sample, 2) responses to a questionnaire asking about ASD diagnosis, and 3) information from state-funded ASD agencies. Children identified as positive through the linking were then invited for diagnostic assessments by the research team.

In total, 60/5,385 children (1.11% of population) with ASD were identified, with 56 of these

5,385 children (1.03% of population) having positive screens on at least one ITC (93% sensitivity) during the first years of life. Some ITCs were positive prior to 12 months. However, in other cases initial screens between 9 to 11 months were negative and did not become positive until later. Some families declined evaluation until two or more screenings were failed. Parental concerns were also tracked independent of ITC status over time. Most parents of children with ASD identified concerns by 24 months of age. The sensitivity of the ITC was above 90 percent at 12 to 14 months through 21 to 41 months, while the proportion of parents indicating concern was less than half under 15 months, increasing to about three-fourths at 21 to 24 months.

As reviewed above in the section on the M-CHAT, an additional fair quality study (Miller, 2011) screened toddlers ages 14 to 30 months with either the ITC, the M-CHAT, or both.[63] The authors did not clearly present the properties of the ITC beyond their combined discussion of performance of the total screening protocol, leaving no information on the PPV of the ITC in isolation.

First Year Inventory (FYI)

Key Summary Points

- Two fair quality studies conducted in the United States and in Israel assessed the FYI and included a total of 611 children screened at 12 months of age.
- Attrition between identification of high risk and diagnostic assessment was over 80 percent.
- No valid performance characteristics are available for this screener.

Description of Measure

The First Year Inventory (FYI) is a 63-item parent report questionnaire designed to assess ASD risk in 12-month-old children. It consists of social-communication and sensory-regulatory domains that sum to form a total risk score.

Detailed Analysis

A fair quality study conducted in the United States[79] examined the use of the FYI in a community sample of 1,305 12-month-old children (Turner-Brown 2014) whose parents completed a mailed copy of the FYI. Over 1,000 (n=1,192) of these families agreed to participate in future research and were re-contacted within 6 months following their child's third birthday. Six hundred and ninety-nine families responded and completed a mailed packet including the Social Responsiveness Scale-Preschool (SRS-P) and Developmental Concerns Questionnaire (DCQ). Of these 699 responders, researchers invited back any responding family whose child was deemed at risk based on one of four criteria: 12-month FYI scores >90th percentile, SRS-P at or above a total score of 60, ASD-related concerns noted on the DCQ, or mild collective concerns noted across measures. Although 153 children met these criteria, only 28/153 (18%) completed diagnostic evaluations; we do not provide performance characteristics due to the extremely high attrition (82%).

Another fair quality pilot study conducted in Israel examined the use of a 24-item version of the FYI, the FYI-Lite (FYI-L-Hebrew), to determine ASD risk in a sample of 583 children at 12-month well-child checks (Ben-Sasson 2013).[58] Fifteen of the 583 (3%) children screened positive using a cut-off score of 0.42 (the U.S. 98th percentile cut-off and Israeli 95th percentile cut-off). Three of them (20%) did not complete follow-up evaluations. Of the remaining 12, 10 (two refused) completed research evaluations consisting of the Mullen Scales and Autism Observation Scale for Infants. Six of these 10 (60%) were referred due to developmental concerns, and one of the 10 (10%) was determined to be at ASD risk. A control group of 12 children also completed evaluations. Researchers then reviewed the medical records of those at-risk children who reached age 24 months during the course of the study. Five of the original 15 screen positive children had medical record information available. Three of these five were noted to have social-communication concerns in the record. In addition to the group who received research evaluations, medical records were reviewed for a subset of children who screened negative on the FYI-L (n=148). Seven of those 148 (5%) had documented social-communication delays in the medical record.

Early Screening of Autistic Traits Questionnaire (ESAT)

Key Summary Points

- Use of two-stage ESAT screening at 14-15 months of age identified only a small number of children with ASD in a large population sample (n=31,724) in one fair quality study conducted in the Netherlands.
- Targeted clinical surveillance and concern identified more children (n=39) with ASD than use of the ESAT (n=18).
- The instrument has not been studied extensively in primarily English speaking populations although many translations of the instrument are available.

Description of Measure

The Early Screening of Autistic Traits (ESAT) tool is a 14-item screening instrument and procedure designed to help identify ASD between 14 and15 months of age in combination with specific developmental surveillance. Children are prescreened with a 4-item version of the ESAT at well-child checks; subsequently, for children screening positive on the 4-item measure, a 14-item version of the questionnaire is completed by a home behavioral professional with parental input. Items on both instruments refer to key areas of social development and play behavior.

Detailed Analysis

One fair quality study conducted in the Netherlands followed the two-step screening process (Dietz 2006).[65,66] Screen 1 used a 4-item version of the ESAT (available in English, Arabic, and Turkish) to screen a random population sample of 31,724 children in well-baby clinics at ages 14-15 months. Children screened positive if their providers in collaboration with parents endorsed 1 of 4 items. Three hundred and seventy of the 31,724 children (1%) screened positive on the 4-item ESAT, of whom 255/370 (69%) completed Screen 2 at approximately 16 months of age. An additional 109 children were identified based on clinician concern. Screen 2 used the

14-item ESAT and was conducted in the home by a psychologist. One hundred children screened positive at Screen 2 by failing at least 3 items. Of these 100, 73 children completed a more comprehensive diagnostic assessment by the study team, including a child psychiatrist. Eighteen of 73 children (25%) were identified with ASD as a result of this process, with 55/73 (75%) receiving other diagnoses (13 intellectual disability, 18 language delay, 25 other behavioral diagnoses). The proportion of children identified as having ASD was substantially lower (0.57 per 1,000) than the known population prevalence, even with high followup (61%). More children with ASD were identified via clinician concern (n=39) than with the ESAT.

Social Attention and Communication Study (SACS)

Key Summary Points

- One fair quality Australian study including 20,770 children screened between 8 and 24 months of age assessed the SACS.
- The prevalence of children identified with the SACS is 0.43 percent. No information is available regarding false negatives.
- The instrument required specialized training of providers in use of the key observational checklists.

Description of Measure

The Social Attention and Communication Study (SACS) measure is an observational tool designed to be completed by maternal and child health nurses conducting well-visits with infants and toddlers. Specifically, nurses complete observational ratings of children on a variety of social and communication developmental milestones at 8, 12, 18, and 24 month well-child visits. Children failing specific combinations of critical items at these specified time points are identified as at-risk for ASD from 12 months onward.

Detailed Analysis

This fair quality, population-based (20,770 children) study included children attending their well-child visits at maternal and child health centers within a 20-km radius of Melbourne University in Victoria, Australia (Barbaro 2011).[73,74] Maternal and child health nurses, who had been trained to administer and score the measure, screened the children. The nurses were trained to look for a pattern of key items associated with ASD at the different age points (8, 12, 18, or 24 months). If children failed a series of these items they were referred on for further evaluation. Two hundred and sixteen of the 20,770 children (1.04%) were referred for a comprehensive developmental assessment (i.e., screened positive on the assessment), and 110/216 children (51%) completed diagnostic follow up. Of these 110 children, 89 (81%) were classified with an ASD, with 20 (18%) children receiving diagnoses of developmental delay or language disorder.

Young Autism and Other Developmental Disorders Checkup Tool (YACHT)

Key Summary Points

- One fair quality study conducted in Japan and including 2,814 toddlers screened at 18 months of age assessed the YACHT.
- Screening with elements of the YACHT as early as 18 months of age identified some cases of ASD within community samples of Japanese children.
- Little information is available about screen negatives.
- The proportion of children identified with elements of the YACHT is 0.39 percent.

Description of Measure

The Young Autism and Other Developmental Disorders Checkup Tool (YACHT) consists of a developmental questionnaire (i.e., motor functioning, communication, social interaction), a caregiver interview regarding pointing and language comprehension, and a specific examination of children asking them to point to identified picture cards.

Detailed Analysis

A fair quality study conducted in Yokohoma, Japan used a children's health surveillance program to attempt to increase the efficacy of early detection (Honda 2009).[72] Public health nurses screened children at their routine 18-month check-up using the Young Autism and other developmental disorders Checkup Tool (YACHT-18) in two stages. Children who screened positive on the YACHT-18 received a followup with a telephone call or home visit as well as individual psychological consultations and a weekly group program for mothers and children. Those needing specialized assessment were seen at a clinic and evaluated by a developmental psychiatrist, a clinical psychologist, and a social worker along with public health nurses. Those with strong suspicion of a developmental disorder were referred to the Yokohama rehabilitation outpatient clinic for diagnostic evaluation (methodology not explained). Those screening negative or children not taking part in the 18-month check-up were assessed again at the 3-year check-up/screening. Those screening positive at the 3-year check-up along with those referred based on community concern (i.e., from kindergartens, nursery schools, other medical clinics, and child guidance clinics) received diagnostic evaluations.

Of 2,814 toddlers screened, 402 screened positive and 2412 screened negative for any developmental disorder at 18 months. Among those screening positive and followed up (number not provided), 19 (7%) were referred for diagnosis (mean age=2 years, 11 months). Four children who screened negative were also referred later (mean age=4 years, 4 months), two because of parental concern and two who screened positive at the 3-year health check. Of the 23 cases screening positive or referred for follow-up, 14/23 (61%) were diagnosed with ASD (5 with autistic disorder and 9 with PDD-NOS). YACHT-18 correctly identified 11/14 cases (79%) of ASD. YACHT-18 missed three actual cases of ASD (22%), including two children with autistic disorder and one with PDD-NOS. One child with language disorder, five with ADHD, and two with intellectual disability received false positive scores on the YACHT-18. Among those

identified by YACHT-18 to be at risk for ASD, 58 percent were actually confirmed by diagnostic evaluation. Of note, the study did not clearly define criteria for screening positive and did not follow the same age criteria for diagnostic evaluation.

Tables 3 and **4** summarize performance characteristics for all studies, and **Table 5** summarizes findings from larger studies with correction for attrition.

Key Question 2a. Do Certain Risk Factors Modify the Performance Characteristics of ASD Screening Tests?

We did not identify any studies that took place in primary care settings that assessed the effect of risk factors such as prematurity or sibling status on performance characteristics of ASD screening tests.

Key Question 2b. Does the Age at Which ASD Screening Is Performed Modify the Performance Characteristics of ASD Screening Tests?

The above reviewed studies of primary care screening did not systematically examine the performance characteristics of early screening tools for children at different ages. However, one study,[70] including a subsample of high- and low-risk children also reported in other M-CHAT studies,[101] attempted to examine screening characteristics of the M-CHAT at different ages, but all within the very young age group of <30 months. The researchers examined outcomes for low-risk children between 17-23 months of age (n=4265, mean age= 18.57 months) and at 24-30 months of age (n=1785, mean age=24.74 months). PPV for children at older ages (0.61) was better than the younger group (0.28). Because this study had already excluded children who had previously been identified as being of concern for developmental delays, the performance characteristics are likely not reflective of what might be seen in the complete population. It also provides no data on screening children at preschool ages versus older ages. Data on false negative were unavailable.

Key Question 2c. Do Other Characteristics of the Child or Family Modify the Performance Characteristics of ASD Screening Tests?

No studies were available to assess whether characteristics of the child and family modify performance characteristics of screening tests.

Key Question 3. What Are the Harms of ASD Screening for the Child and Family?

No studies assessed or addressed harms of screening.

Key Question 4. What Is the Effect of Interventions Targeting Young Children (in Preschool and Elementary School) on the Outcomes of Core ASD Symptoms, Cognitive and Intellectual Functioning, Language and Communication Skill Development, Challenging Behavior, Adaptive Behavior, Educational Placement/Achievement, and Quality of Life for the Child and Family?

Treatments for young children in the target age group for routine screening for ASD are primarily behavioral interventions, particularly early intensive behavioral and developmental interventions, which may include approaches incorporating applied behavior analysis (ABA) principles and/or approaches incorporating parent training components, and play/interaction-based interventions. **Table 6** outlines key cognitive and language outcome measures, and we present the results of early intervention studies below and in **Tables 7-14** and summarize them in the discussion. Because the most evidence is available for the effects of ABA-based interventions on cognitive and language outcomes, we describe these in most detail below. Data on other outcomes are available in tables in **Appendix F**. Finally, we very briefly summarize play-based therapies, which have a substantially weaker body of evidence.

Approaches for Younger Children

Early Intensive Behavioral and Developmental Interventions

Key Summary Points

- Forty-two unique studies addressed early intensive behavioral and development interventions for children with ASD. These included 26 RCTs (nine good and 17 fair quality), five non-randomized trials (one good and four fair quality), 10 prospective studies (two good and eight fair quality), and one retrospective cohort study (fair quality). None of the studies focused on screen detected children.
- All studies compared a minimum of two treatment groups. No study included a control group that was not receiving some type of intervention (including school enrollment or eclectic community-based therapies, such as medication or occupational therapy), although some limited the number of behaviorally based treatment hours that control participants could receive.
- Data are inadequate to predict which children are most likely to benefit from early intervention, although benefits achieved differ by child characteristics and interventions

offered.
- No studies reported harms.

Direct provision to child:
- Seventeen studies (two good and two fair quality RCTs; one good and four fair quality non-randomized trials; eight fair quality cohort studies) assessed studies of early intervention involving direct provision of treatment to the child.
- Some young children receiving high intensity child-focused over 8 months to 2 years demonstrated statistically significant improvements in cognitive and language outcomes with these approaches compared to eclectic community based interventions.

Parent training:
- Thirteen studies included parent training approaches (4 good and 6 fair quality RCTs; 2 good and one fair quality cohort studies).
- Studies with parent training components reported some improvements in language with inconsistent results for other outcomes.

Play/interaction-focused:
- Twelve unique RCTs (three good and nine fair quality) addressed play or interaction-based approaches.
- Most studies targeted aspects of joint attention and reported statistically significant gains in joint attention skills in treatment groups compared with controls, typically over a short duration (8 to 12 weeks). Children in both treatment and comparison groups typically received early intervention in addition to the targeted intervention.
- This body of evidence for the effects of play-based interventions is limited due to very small studies, few total participants across all studies (n=483), and outcomes too heterogeneous to pool.

We identified 42 studies (reported in 55 publications) of good (n=12) and fair (n=30) quality addressing early intensive behavioral and developmental interventions.[103-157] Early intensive behavioral and developmental interventions are the most commonly recommended approaches for young children. These approaches can either be "manualized" in that a specific manual is developed and published and followed carefully, or more general, in which developmental approaches are applied in different ways. These interventions are based on the principles of ABA. ABA is an umbrella term describing principles and techniques used in the assessment, treatment and prevention of challenging behaviors and the promotion of new desired behaviors. The goal of ABA is to teach new skills, promote generalization of these skills, and reduce challenging behaviors with systematic reinforcement. The principles and techniques of ABA existed for decades prior to specific application and study within ASD, and ABA provides the basis for many early intensive behavioral and developmental approaches.

The principles of ABA may be applied in models of care in which the provider works directly with the child, including, for example the Early Start Denver Model; University of California/Lovaas-based interventions; preschool-delivered general ABA-based early intervention approaches, and the Learning Experiences and Alternative Program for Preschoolers and their Parents (LEAP) model. A total of 17 studies (**Table 7**) fall into this category (two good

and two fair quality RCTs, one good and four fair quality non-randomized trials, seven fair quality prospective cohort studies, and one fair quality retrospective cohort study).

Another model of care is early intensive interventions with parent-training components (**Table 8**), and these include parent-led ESDM, general home-based or clinic-based early intensive intervention, Pivotal Response Training (PRT), the Assessment Evaluation and Programming System for Infants and Children (AEPS), the Developmental, Individual Difference, Relationship-based (DIR)/Floortime model, parent training in communication responsiveness, and More than Words. Thirteen studies addressed these interventions (four good and six fair quality RCTs and two good and one fair quality prospective cohort studies).

Finally, 12 studies took an approach that was play-based and focused on interaction between child and parent or child and provider.

These interventions are generally focused on improving cognitive and language outcomes, and the results are presented below by target outcome. Measures of adaptive behavior and ASD severity are also available for these interventions and are provided in **Appendix F**.

Studies of Direct Provision to Child Approaches: Cognitive Outcomes

Fifteen of 17 studies of ABA-based interventions provided data on cognitive outcomes associated with treatment, of which four were RCTs and 11 were non-randomized trials or cohort studies. In 11 of the 15 studies, the comparator was "eclectic treatment," which means that parents were obtaining treatment in the community for their children, but the investigators did not control what that treatment was. Seven of these 11 studies with eclectic comparison groups evaluated cognitive outcomes, and outcomes were improved in the early intervention group compared with eclectic in seven. Across all comparison groups (eclectic or other comparison group), the ABA-based intensive intervention resulted in statistically significantly greater cognitive improvements than did the comparator in 10 of 15 studies evaluating cognitive changes.

UCLA/Lovaas approaches. The University of California, Los Angeles (UCLA)/Lovaas approach is a well-known treatment that is most commonly associated with the ABA method. UCLA/Lovaas-based approaches draw on ABA-derived principles and typically include high intensity (>20 hours per week) one-on-one instruction with primary emphasis on discrete trial techniques. Discrete trial training introduces a stimulus (instruction/cue) to which a child may respond. Responses may be reinforced/rewarded, and the trial of stimulus-potential response-reward is repeated to develop mastery. These approaches also emphasize incidental teaching and generalization of skills to the home or other settings. Approaches that are considered UCLA/Lovaas may have some variation but all base their treatment on the UCLA/Lovaas manual.

Eight studies assessing UCLA/Lovaas-based models provide data to assess cognitive outcomes. Four of these eight studies compared UCLA/Lovaas models to eclectic or community-based treatments, two compared clinic-directed interventions to parent-managed interventions (parents identifying and directing ABA therapists), one compared a therapist-led to a parent-led approach,

and one compared high intensity to low intensity intervention.

Overall, children in the treatment arms were more likely to improve significantly than children in the comparators. In three of the four studies that compared the treatment to eclectic treatment, cognitive measures were significantly improved in the UCLA/Lovaas groups compared with eclectic controls. In the fourth, there was no difference. The two studies comparing clinic-directed UCLA/Lovaas to parent-managed interventions demonstrated no significant difference in cognitive outcomes. The one study comparing a therapist-led UCLA/Lovaas model with parent-led reported significantly improved outcomes for the therapist-led arm. The study comparing high to low intensity found greater effects associated with higher intensity intervention. Specific treatment effects varied and are described below.

UCLA/Lovaas vs. eclectic approaches. Three of the studies comparing the UCLA/Lovaas approach to eclectic approaches were nonrandomized trials, and the fourth was a prospective cohort study. Overall, cognitive outcomes were significantly improved in the UCLA/Lovaas groups compared with the control in three studies and not significantly different between groups in the fourth.

In one fair quality non-randomized trial (Peters-Scheffer 2011) comparing a UCLA/Lovaas model implemented in a specialized ASD preschool with a specialized ASD preschool alone,[114] cognitive outcomes measured on the Bayley Scales of Infant Development were significantly improved ($p<0.01$) in the UCLA/Lovaas group compared with control (treatment group baseline: 47.00 ± 10.33, followup: 55.83 ± 14.94; control group baseline: 45.73 ± 15.99 followup: 43.73 ± 16.74). All children in this study had concomitant intellectual disability in addition to an ASD diagnosis.

A second fair quality non-randomized trial (Eikeseth 2007) compared an intensive, school-based UCLA/Lovaas program with school-based eclectic treatment that included elements of multiple interventions including the Treatment and Education of Autistic and Related Communication-Handicapped Children (TEACCH) model and occupational therapy.[117,118] Followup scores for individual cognitive outcomes (measured on the Bayley Scales, Wechsler Preschool and Primary Scale of Intelligence-Revised [WPPSI-R], or Wechsler Intelligence Scales for Children-Revised [WISC-R]) did not differ significantly between groups; however, gains from baseline to followup were significantly greater in the UCLA/Lovaas group compared with the eclectic group (mean cognitive score change of 17.15±10.97 vs. 4.33±7.55, $p<0.01$). Children in the treatment group were significantly more likely to have cognitive scores in the average range at followup than the eclectic group (7/13 children vs. 2/12, $p<0.05$). At followup of all participants roughly 31 months after treatment onset, cognitive gains continued to be significantly greater ($p<0.05$) in the UCLA/Lovaas group compared with the eclectic arm (mean gain of 25 vs. 7 points gained). Children in both groups continued with either UCLA/Lovaas therapy or eclectic therapy, both at reduced intensity, during the followup interval.

One fair quality prospective cohort study (Cohen 2006) compared intensive UCLA/Lovaas-based intervention delivered in the school and home to eclectic therapies available in the community and delivered primarily in the school.[119] Cognitive skills measured on the Bayley Scales or WPPSI increased significantly in the UCLA/Lovaas group compared with the eclectic

comparators (p<0.05), though mean cognitive scores increased in both groups (gain of 25 points in treatment group and 14 points in control). Twelve of the 21 children in the intervention group and 7 of the 21 in the comparison group had IQs in the average range at followup (p=NS).

Finally, one good quality non-randomized trial (Peters-Scheffer 2013) compared low intensity (4-10 hours/week) UCLA/Lovaas-based intervention added to preschool programming for children with ASD *and* intellectual disability with eclectic, community-based intervention, including specialized preschool.[121] Cognitive outcomes measured on the Mullen Scales of Early Learning improved in the treatment group at the 12-month but not the 24-month assessment (effect size=0.40) and did not improve at either time in the control group. Between group differences were not significant.

Clinic-directed vs. parent-managed UCLA/Lovaas approaches. Two studies compared clinic-directed approaches to parent-managed care. In both studies, all children improved their scores on cognitive outcomes, but neither whether the treatment was directed by the clinic or managed by the parents was not associated with a difference.

The first, a good quality RCT (Sallows 2005), compared intensive (mean of approximately 37-39 hours/week) clinic-directed UCLA/Lovaas therapy to UCLA/Lovaas-based therapy (approximately 31-32 hours/week) with therapists managed by parents, who set the number of therapy hours.[122] Cognitive outcomes as measured on various age-appropriate tools (Bayley Scales of Infant Development-II, WPPSI-R, WISC-III) did not differ significantly between groups at baseline or followup. In analyses combining groups, mean full scale IQ increased from a mean of 51 to 76 points, and children improved significantly on full scale IQ, verbal IQ, and performance IQ from baseline in both groups (all p values <0.01).

An additional fair quality non-randomized trial compared a clinic-directed, home-based intensive UCLA/Lovaas-based model to a parent-managed model in which parents identified and managed tutors.[115,116] These tutors also received supervision from the same consultants involved in the clinic-directed model. Cognitive outcomes as measured on the Bayley Scales of Infant Development (for younger children) or the WPPSI-R improved from baseline to followup in both groups with no significant group differences. In analyses combining groups, mean scores increased 16 points from baseline with half of children gaining 15 or more points (range=15 to 52), 39 percent gaining 1 to 14 points, 2 percent maintaining the baseline cognitive scores, and 9 percent losing 4 to 18 points.

Clinic-conducted vs. parent-conducted UCLA/Lovaas approaches. One, fair quality RCT (Smith 2000) compared intensive, therapist-led UCLA/Lovaas-based intervention with intensive parent-led UCLA/Lovaas-based intervention to assess the ability of parents to provide the intensive approach relative to professional therapists. Cognitive outcomes on the Stanford-Binet Intelligence Scale or Bayley Scales of Infant Development were significantly improved in the therapist-led group compared with parent-led (therapist-led group intake IQ=50.53±11.18, followup IQ=66.49±24.08; parent-led group intake IQ=50.69±13.88, followup IQ=49.67±19.74, p<0.05).[120]

High intensity vs. low intensity UCLA/Lovaas approaches. Finally, one fair quality non-

randomized trial (Reed 2007) compared three models of high intensity (>20 hours/week) ABA-based interventions to generic, low intensity (<20 hours/week) early intervention.[123] The high intensity interventions included UCLA/Lovaas-based intervention; a verbal behavior model that used discrete trial training to target verbal responding and the ability to use language to request and name objects/events; and the Complete Application of Behavior Analysis to Schools (CABAS) approach. CABAS focuses on training teachers in multiple ABA techniques including incidental teaching and reinforcement and involves maximizing child-targeted "learn units" or opportunities to learn. The study assessed cognitive outcomes using the Psycho-educational Profile-Revised and British Ability Scale. In comparisons of all high intensity interventions vs. low intensity, children in high intensity programs improved significantly more in intellectual functioning and educational functioning (p values <0.01) than did children in low intensity programs. In comparisons among high intensity interventions, children in the UCLA/Lovaas and CABAS groups had the greatest gains in intellectual functioning ($p<0.05$). Change scores for educational functioning were significantly higher for the CABAS group (effect size=3.74, $p<0.01$) than the other high intensity groups.

Early Start Denver Model (ESDM). ESDM is an approach directed to young children that integrates ABA-based early intervention approaches with developmental and relationship-based care, with specific activities involving interpersonal communication, engagement with real-life materials and activities, and parental use of the strategies in the home. ESDM is typically provided for approximately 30 hours/week and may be delivered in a child's home or clinic- or school-based settings by trained therapists and parents.

The ESDM approach has been assessed in one good quality RCT and a follow up study, with changes in cognitive scores a primary outcome.[103,104] Cognitive skills measured on the Mullen Scales of Early Learning's Early Learning Composite improved significantly in children in the ESDM group compared with children receiving eclectic, community-based interventions (mean gain of 15.4 vs. 4.4 points, respectively, $p=0.018$). At a 2-year followup, cognitive scores for children in the ESDM group had improved a mean of 17.6 points from baseline compared with a 7-point gain in the control group ($p=0.044$). The average starting cognitive score was 61.0±9.2 for children in the intervention group and 59.4±8.6 for children in the comparison group. The standard score on the Mullen Early Learning composite is 100±15.

LEAP. Another ABA-based approach is the LEAP model, which integrates multiple interventions for children with ASD with modeling and social skills development mediated by typically developing peers in a preschool setting. The model incorporates elements of Pivotal Response Training, the Picture Exchange Communication System (PECS), positive behavior support, and naturalistic, incidental teaching, in which teachers use classroom interactions and a child's interest to guide teaching.

One fair quality RCT (Strain 2011) evaluated the LEAP model and reported cognitive outcomes.[112] Compared with children in preschools with teachers who received LEAP manuals and no specific instruction, children in preschools with teachers who received LEAP coaching and training gained significantly on cognitive measures (mean gain of 8.9 points vs. loss of 1.8 points on the Mullen Early Learning Composite, $p<0.01$).

Preschool-delivered models. Preschool-delivered general ABA-based models combine elements of ABA techniques such as one-on-one instruction using discrete trial techniques and incidental teaching and typically deliver >20 hours/week of intervention. Approaches may draw from numerous ASD treatment types but focus primarily on ABA-based techniques.

Five of six studies assessing these approaches in preschools reported cognitive outcomes, and all five of these compared ABA-based intervention delivered in a preschool environment to eclectic therapies, some of which were school-based. In four of the five studies, cognitive outcomes were significantly improved for the ABA groups vs. the control arms; group differences were not significant in one study. We briefly summarize these studies below. Four of the five were fair quality prospective cohort studies, and one was a fair quality retrospective cohort.

The first prospective cohort study (Eldevik 2012) compared outcomes in children receiving ABA-based intervention in mainstream preschools to treatment-as-usual and measured cognitive outcomes with the Bayley Scales of Infant Development-II or III, Stanford-Binet-IV or V, or WPPSI-R.[107] The ABA group had significantly improved intellectual functioning (p=0.004) compared with the control group (mean change in score of 15.1±14.9 vs. 0.5±9.5, effect size = 1.03, 95% CI: 0.34 to 1.72). A second fair quality prospective cohort study (Howard 2005) compared preschool-based ABA intervention with eclectic treatments received in public school classrooms and with non-intensive early intervention classrooms.[111] Cognitive outcomes as measured primarily on the WPPSI-R were significantly better in the ABA group compared with the other groups (difference in mean followup score for ABA group and mean of other groups combined=24.42, p<0.01). Average scores for the ABA group were 58.54 at baseline and 89.88 at follow-up, compared to a comparison group mean of 68.81 at follow-up. On the WPPSI-R, children who score at 68 are in the bottom 5% of same-aged test-takers. In contrast, children who score at 89 do as well or better than 27% of same-aged test-takers, reflecting children's significantly improved ability to engage in testing and display their skills.

A third fair quality prospective cohort study (Zachor 2007) comparing ABA-based preschool intervention with eclectic developmental intervention reported cognitive outcomes measured on the Bayley Scale of Infant Development-II and Stanford-Binet-IV scales for both groups combined.[110] Cognitive outcomes were significantly improved in the ABA group compared with eclectic at followup.

A fair quality retrospective cohort study (Flanagan 2012) compared outcomes for children enrolled in ABA-based intervention in preschool to outcomes in children on the program waiting list, who were receiving eclectic therapies including speech therapy, specialized or regular daycare, and low-intensity early intervention.[109,141] Cognitive skills were assessed using either the Mullen Scales of Early Learning, the Stanford-Binet-IV, or the WPPSI-III. Cognitive scores were higher in the ABA group compared with control at followup (mean 55.80±26.97 vs. 39.50 ±18.93, p=0.002, effect size=0.83). At followup, 18 percent of the 61 children in the ABA group vs. 3.3 percent of the 61 in the waitlist group had IQ scores in the average range at followup (p=0.008). Finally, a fourth fair quality prospective cohort study (Itzchak 2011) comparing ABA-based intervention and eclectic-developmental interventions reported no significant group differences in followup cognitive outcomes measured on the Mullen Scales of Early Learning.[105,106]

Studies of Direct Provision to Child Approaches: Language Outcomes

The other most commonly assessed outcome is language skills. ASD has a significant communication component, and children may present with difficulties in both expressive and receptive language. The ability to communicate with their children is often a parent-centered outcome. Several of the studies described above that reported on cognitive outcomes also reported on shifts in language. Ten additional studies provided data on language outcomes, but not on cognitive outcomes. Overall, 16 studies of ABA-based early intervention reported language outcomes, with more significant improvements in the ABA arms compared with control arms in 10 studies, and no significant between group differences in six studies.

UCLA/Lovaas-based approaches. Seven of the eight studies also described above assessing UCLA/Lovaas-based models reported language outcomes. Two of the four studies that compared UCLA/Lovaas models to eclectic or community-based treatments reported significantly greater improvements in the treatment group, while no differences were observed in the other two. Two studies compared clinic-directed UCLA/Lovaas to parent-managed interventions, in which parents identified and managed ABA therapists. Language outcomes were not significantly different between groups in either of these studies. One study compared therapist-led and parent-led UCLA/Lovaas models and reported significantly improved language outcomes in the therapist-led arm.

UCLA/Lovaas vs. eclectic approaches. In one fair quality non-randomized trial (Peters-Scheffer 2010) comparing a UCLA/Lovaas model implemented in a specialized ASD preschool with a specialized ASD preschool alone,[114] language outcomes measured on the VABS-II Communication domain improved in both groups over time, with significantly greater gains (p=0.02) in the UCLA/Lovaas group compared with control (baseline communication score in months for the treatment group: 26.92±12.12, followup: 39.42±15.39; baseline score in control group: 25.00±10.00, followup: 29.95±13.39).

One fair quality non-randomized trial (Eikeseth 2007) compared an intensive, school-based UCLA/Lovaas program with school-based eclectic treatment that included elements of multiple interventions including TEACCH and occupational therapy.[117,118] Followup scores for individual language outcomes (measured on the Reynell Developmental Language Scales and VABS-II Communication domain) did not differ significantly between groups; however, gains from baseline to followup were significantly greater in the UCLA/Lovaas group compared with the eclectic group (change on Reynell Developmental Language Scales total score of 27.00±20.41 vs. 1.08±17.07, p<0.01; change on VABS-II communication: 15.69±16.89 vs. -1.58±7.81, p<0.01). At followup of all participants roughly 31 months after treatment onset, VABS-II communication scores (Reynell Developmental Language Scales not measured at long term followup) were significantly higher in the UCLA/Lovaas arm vs. control (mean 78.5±22.3 vs. 56.0±16.3, p<0.01). Children in both groups continued with either UCLA/Lovaas therapy or eclectic therapy, both at reduced intensity, during the followup interval.

One fair quality prospective cohort study (Cohen 2006) compared intensive UCLA/Lovaas-based intervention delivered in the school and home to eclectic therapies available in the community and delivered primarily in the school.[119] Language outcomes measured on the Reynell

Developmental Language Scales were not significantly different between groups at followup. Eight of the 21 children in the intervention group and 4 of the 21 in the comparison group had language comprehension scores in the average range at followup (p=NS). Nine children in the intervention group and 6 in the comparison arms also had expressive language scores in the average range (p=NS).

One good quality non-randomized trial (Peters-Scheffer 2013) compared low intensity (4-10 hours/week) UCLA/Lovaas-based intervention added to preschool programming for children with ASD and intellectual disability with eclectic, community-based intervention, including specialized preschool.[121] Language outcomes were measured using the Early Social Communication Scales, Peabody Picture Vocabulary Test (which measured expressive language), VABS-II Communication domain, and Schlichting Test for Language Production. Receptive, but not expressive, language improved significantly in the UCLA/Lovaas group compared with control (p=0.04, effect size=1.22), and group differences were not significant for any subscale of the early social communication measure. The UCLA group also improved significantly on the VABS-II communication domain compared with the control group (p<0.004, effect size=1.41).

Clinic-directed vs. parent-managed UCLA/Lovaas approaches. One fair quality non-randomized trial (Hayward 2009) compared a clinic-directed, home-based intensive UCLA/Lovaas model to a parent-managed model in which parents identified and managed tutors.[115,116] These tutors also received supervision from the same consultants involved in the clinic-directed model. Language outcomes as measured on the Reynell Developmental Language Scales improved from baseline to followup in both groups with no significant between-group differences. In analyses combining groups, mean language comprehension age increased to 27.5 months and mean expressive language age increased to 26.95 months from means of roughly 21 months in each group at baseline.

A good quality RCT (Sallows 2005) compared intensive (mean of approximately 37-39 hours/week) clinic-directed UCLA/Lovaas therapy to UCLA/Lovaas-based therapy (approximately 31-32 hours/week) with therapists managed by parents, who set the number of therapy hours.[122] Language outcomes as measured on various tools as age-appropriate (Reynell Developmental Language Scales, Clinical Evaluation of Language Fundamentals, VABS communication domain) did not differ significantly between groups at baseline or followup. In analyses combining groups, children improved significantly on receptive language and the VABS communication domain from baseline (all p values <0.01).

Clinic-conducted vs. parent-conducted UCLA/Lovaas approaches. In a fair quality RCT (Smith 2000) comparing intensive, therapist-led UCLA/Lovaas-based intervention with intensive parent-lead UCLA/Lovaas-based intervention, overall language outcomes on the Reynell Developmental Language Scales were significantly improved (p<0.05) in the therapist-led group compared with parent-led (therapist-led group total Reynell score at intake=28.60±4.07, followup=87.40±46.21; parent-led group intake=30.00±6.34, followup=61.33±31.88).[120] Scores on the Reynell Developmental Language Scales Expressive Language and Language Comprehension subscales did not differ significantly between groups nor did scores on the VABS communication domain.

ESDM. The ESDM RCT described above also reported language outcomes.[103,104] Receptive and expressive language measured on the Mullen Scales of Early Learning improved significantly in the ESDM group compared with the comparison arm (ESDM mean gains: 18.9 and 12.1 points vs. 10.2 and 4.0 points in comparison arm) at two years post-treatment (p values <0.05). In the ESDM group children at an average of age of 23.9 months had an Early Learning Composite (SS Mean=100, SD=15) of 59.4 (1st percentile). This corresponded to domain scores (T scores Mean=50, SD=10) of 21.1 (1st percentile) regarding language understanding (Receptive Language); 24.5 regarding language use (1st percentile; Expressive Language); 33.2 (4th percentile) regarding nonverbal problem solving (Visual Reception), and 33.9 (4th percentile) regarding fine motor skills (Fine Motor). At 52.4 months the ESDM group evidenced an Early Learning Composite (SS Mean=100, SD=15) of 78.6 (7th to 8th percentile). This corresponded to domain scores (T scores Mean=50, SD=10) of 40.1 (16th percentile) regarding language understanding (Receptive Language); 38.6 regarding language use (12th percentile; Expressive Language); 41 (18th percentile) regarding nonverbal problem solving (Visual Reception), and 33.5 (4th percentile) regarding fine motor skills (Fine Motor). Children in the other treatment arm (community treatment) did not exhibit statistically significant difference on their cognitive profile.

At an average of 52.4 months of age at outcome assessment (e.g., 28.5 months later) children in the ESDM (treatment) group had new developmental age equivalents of approximately 44 months, Receptive Language; 43 months, Expressive Language; 46 months, Visual Reception; and 42 months, Fine Motor. The community treatment group, at an average of 52.1 months at outcome (e.g., 29.1 months later) had new developmental age equivalents of approximately 37 months, Receptive Language; 36 months, Expressive Language; 41 months, Visual Reception; and 40 months, Fine Motor. Both groups improved, with the active treatment groups improving significantly more. A description of the potential impact of this level of change is provided in the Discussion.

LEAP. Two studies of LEAP assessed language outcomes with mixed results (improved outcomes in the full LEAP training arm compared with control in one and no significant differences among groups in another). One of these is also described above and included cognitive outcomes.

The fair quality RCT described above (Strain 2011)[112] reported that compared with children in preschools with teachers who received LEAP manuals and no specific instruction, children in preschools with teachers who received LEAP coaching and training gained significantly on language as measured by the Preschool Language Scales-Fourth edition (mean gain of 18.5 points vs. 9.4 points, p<0.01).

In addition, another fair quality prospective cohort study (Boyd 2013) assessed language outcomes using a composite score derived from the Preschool Language Scales, 4th edition, Mullen Scales of Early Learning Expressive and Receptive language scales, and Vineland Adaptive Behavior Scales, 2nd edition (VABS-II) Expressive and Receptive Communication scales.[113] The study compared outcomes in children with ASD enrolled in three types of preschools: 1) those using the LEAP model 2) those using the TEACCH model, which uses elements such as visual schedules and structured learning environments to promote children's

engagement and learning, and 3) those not using any specific model. Communication composite scores improved significantly in all three groups from baseline (all p values ≤0.05), but between group differences were not significant.

Preschool-delivered models. All six studies that assessed ABA models in preschools reported language outcomes, and all of these compared ABA-based intervention delivered in a preschool environment to eclectic intervention. In five of the six studies, language outcomes were significantly improved for the ABA group while group differences were not significant in one study. We briefly summarize these studies below.

Four studies measured outcomes on the VABS communication domain.[105-109] All four were cohort studies comparing ABA-based intervention to eclectic comparators. In three of the four studies, the treatment group demonstrated more improvement in language than the comparison. In one, the mean change score for children receiving the treatment was 8.6±14.6 compared to 0±12.6 in the comparison group. The study reported p-value for difference was 0.034.[107] In the second, the difference was statistically significant at $p<0.05$, and an effect size on the communication subscale was 1.08.[108] A third compared outcomes for children enrolled in ABA-based intervention in preschool to outcomes in children on the program waiting list, who were receiving eclectic therapies including speech therapy, specialized or regular daycare, and low-intensity early intervention.[109] Language skills were significantly improved in the ABA group vs. control (mean 46.60±29.92 vs. 30.33±16.98, p=0.006, effect size=0.56) at followup. Nonetheless, another similar study demonstrated no difference.[105,106]

One fair quality prospective cohort study (Howard 2005) compared preschool-based ABA intervention with eclectic treatments received in public school classroom and with non-intensive early intervention classrooms.[111] Language outcomes in this study were measured primarily on the Reynell Developmental Scales. Language outcomes were significantly improved for the ABA group vs. the other groups (receptive language difference in mean followup score for ABA group and mean of other groups combined=21.73, $p<0.01$; expressive language difference in mean followup score for ABA group and mean of other groups combined=23.21, $p<0.01$). These significant improvements in receptive and expressive language reflect an improved ability for children to follow instructions, identify objects, and use words and gestures to functionally communicate. This in turn allows therapists to build even more broadly on children's language abilities to teach more complex communication skills.

Finally, an additional fair quality prospective cohort (Zachor 2007) comparing ABA-based preschool intervention with eclectic developmental intervention reported language outcomes measured on the ADOS language and communication subscale.[110] Outcomes were significantly improved for the ABA group compared with the control (mean 7.2±4.1 vs. 9.7±3.0, $p<0.01$).

Studies of Parent Training Approaches: Cognitive Outcomes

Another group of studies (n=13) focused on the application of ABA-based methods via the parents. In these interventions, the parents are trained in an ABA approach and implement it at home. Five of these 13 studies addressed cognitive outcomes. Outcomes did not differ significantly between groups in four studies, and cognitive outcomes were significantly improved in the treatment group compared with the eclectic control group in one study.

Parent-delivered ESDM. As described above, ESDM integrates ABA-based early intervention approaches with developmental and relationship-based approaches that involve interpersonal communication, engagement with real-life materials and activities, and parental use of strategies in the home. ESDM may be delivered in a child's home by trained therapists or parents. One fair quality RCT (Rogers 2012) compared a parent-delivered ESDM model, in which parents received therapist-led training and coaching in using ESDM techniques with their children, with community-based treatments.[129,130] At followup, cognitive outcomes as measured on the Mullen Scales of Early Learning's Early Learning Composite did not differ significantly between groups though both groups improved from baseline (effect size in ESDM group=0.44, control=0.37).

Assessment Evaluation and Programming System for Infants and Children (AEPS). The AEPS is a developmental curriculum that sets targeted goals for children based on their developmental profile. It promotes engaging and sharing attention and imitation of skills and actions. This approach was studied in one good quality RCT (Landa 2011) that compared an intervention combining AEPS, parent training sessions, and additional therapist-led intervention to promote joint attention, imitation of social actions, and sharing positive affect with the AEPS plus parent training sessions alone.[124,125] Non-verbal IQ as rated on the Mullen Scales of Early Learning did not differ significantly between groups at followup. At a followup of the enhanced intervention group at roughly 72 months of age (mean age 27 months at baseline), overall IQ was significantly improved from baseline (mean change=21.4±22.9, effect size=1.02, $p<0.001$).

Clinic and/or home-based early intervention. These models incorporate elements of ABA, discrete trial training, one-on-one training, and naturalistic or incidental teaching in comprehensive programs based in specialized centers and/or offered in the home. Programs train parents in techniques to use in the child's natural environments (such as the home), and intervention is provided in centers and the home by therapists and parents for 20 to 40 hours/week.

Two studies that used clinic and/or home-based models compared with different comparison groups provided data on cognitive outcomes. In one good quality prospective cohort study (Strauss 2012), children in the parent training arm showed a greater increase in mental development compared with children receiving eclectic therapies (change on the Griffiths Mental Development Scales for Ages 2 to 8=4.639 in the parent training group vs. 0.332 in eclectic group, p=NR).[126,127] In addition, a fair quality prospective cohort study (Reed 2012) compared four early intervention models: home-based early intervention using ABA-principles, specialized daycare, a low-intensity special education one-on-one teaching program, and home-based therapist-delivered intervention plus parent training.[132] Groups did not differ on measures of intellectual functioning (measured on the Psycho-Educational Profile-Revised) at followup. The ABA group did have significantly better scores on measures of educational functioning (measured on the British Ability Scale) compared with the other groups ($p<0.05$).

Parent training in communication responsiveness. Intervention models focusing on parental responsivity target parent behaviors and responses in order to promote the development of language and social skills. Such models emphasize shared engagement with the child and parental sensitivity to the child's communication attempts. Models typically promote

development of early language skills such as joint attention and build upon skills as the child's social communication abilities increase. A fair quality RCT (Drew 2002) assessed a parent training intervention to promote early communication skills compared with services available in the community and reported no significant differences between groups at followup on nonverbal IQ (Griffiths Scale of Infant Development).[136]

Studies of Parent Training Approaches: Language Outcomes

Twelve of 13 studies with parent training components reported language outcomes. In three studies, language outcomes were significantly improved in the treatment arms compared with control, and groups did not differ significantly in nine studies. In two of these nine studies, language outcomes as rated by external assessors did not differ between groups, but parent-rated language outcomes were significantly improved in the treatment group compared with the control group.

Parent-delivered ESDM. The fair quality RCT described above also provided outcomes data on language acquisition.[129,130] Language outcomes measured on the MacArthur-Bates Communicative Development Inventory and VABS communication domain did not differ significantly between groups at followup.

Assessment Evaluation and Programming System for Infants and Children (AEPS). The AEPS study described above also provided data on communication outcomes.[124,125] Socially engaged imitation (imitation paired with eye contact with examiner) measured by coding videotapes of child and therapist interaction was significantly higher in the enhanced intervention group compared with controls, but initiations of joint attention and shared positive affect did not differ significantly between groups. Expressive language outcomes measured on the Mullen Scale did not differ significantly between groups at followup. At a followup of only the enhanced intervention group at roughly 72 months of age (mean age 27 months at baseline), VABS communication domain scores were significantly improved from baseline (mean change=12.7± 19.4, effect size=0.81, $p<0.001$).

Clinic and/or home-based early intervention. Three studies used clinic and/or home-based models compared with different comparison groups, and two of these reported language outcomes. One study compared a clinic and home-based program to eclectic therapies and reported improved language outcomes in the center/home-based arm compared with control. One study compared a home-based program that incorporated parent training with a center-based variant of the program and with a wait list control group and reported significantly higher scores in the center-based group compared with home-based for one language measure.

One good quality RCT (Roberts 2011) compared a home-based program that incorporated parent training with a center-based variant of the program incorporating small-group intervention and parent training and with a wait list control group.[131] Children in the wait list group were receiving community-based interventions (interventions not specified). Children in the center-based group improved significantly on Reynell Developmental Language Scale comprehension standard scores compared with the home-based arm (mean difference=7.3, 95% CI: 0.7 to 13.9, p=0.03), but scores in the expressive language subscale did not differ among groups, nor did

scores on the VABS communication domain. Scores on the Pragmatics Profile of Everyday Communication also did not differ significantly among groups.

In one good quality prospective cohort study (Strauss 2012), children in the parent training arm showed a greater increase in language comprehension measured on the MacArthur Communication Development Inventories (change of 6.460 vs. 3.885, p=NR) and in language production (change of 3.410 vs. 1.69, p=NR) compared with children receiving eclectic therapies.[126,127]

Parent training in communication responsiveness. Intervention models focusing on parental responsivity target parent behaviors and responses in order to promote the development of language and social skills. Such models emphasize shared engagement with the child and parental sensitivity to the child's communication attempts. Models typically promote development of early language skills such as joint attention and build upon skills as the child's social communication abilities increase.

Five studies assessed parent training in communication responsiveness and reported language outcomes. Four of these studies compared the parent training arm to eclectic/community-based interventions or treatment-as-usual and one compared a parent training arm with instruction provided by a therapist to parent training with instruction provided via a video. Language outcomes at followup were significantly improved in the active parent training arm compared with control in two studies and did not differ significantly between groups in two. In one study language outcomes evaluated by assessors were not significantly different between groups, but parent ratings favored the treatment group vs. control. We summarize studies briefly below.

One good quality RCT (Aldred 2011) compared a parent training intervention focusing on social communication and parental responses to communication with routine, community-based care.[137,138] Expressive language measured on the MacArthur-Bates Communicative Development Inventory was significantly improved in the treatment arm compared with control (mean scores at followup 199.4 vs. 33.1, $p<0.001$). Language comprehension and VABS communication domain scores did not differ significantly between groups.

One fair quality RCT (Green 2010) comparing a parent responsivity-focused intervention with eclectic treatment as usual reported no group differences on blinded assessor-rated language measures (Preschool Language Scale, coded parent-child interactions), but in parent-rated measures, the treatment group showed significant improvement on MacArthur-Bates Communicative Development Inventory Receptive and Expressive language scales and the Communication and Symbolic Behavior Scales Developmental Profile.[135] Teacher-rated measures (VABS communication domain) did not differ significantly between groups.

A good quality prospective cohort (Keen 2010) compared a professionally supported parent training intervention with a self-directed video-led program.[139] Both programs focused on parental responsivity to child communication and integrating intervention into family routines. Followup scores on Communication and Symbolic Behavior Scales-Developmental Profile social communication measures were significantly more improved in the professionally supported arm compared with the self-directed arm on parent-rated measures but not on

observer-rated measures.

A fair quality RCT (Drew 2002) assessed a parent training intervention to promote early communication skills compared with services available in the community and reported no significant differences between groups at followup on language measures (MacArthur Inventory), though more children in the parent training group compared with control moved from being nonverbal to having single word or phrase speech (n=7 vs. 3, $p<0.05$).[136] A fair quality RCT (Oosterling 2010) replicated this intervention model in an ASD preschool setting, comparing parent training plus ASD preschool to ASD preschool alone.[133] No language outcomes (measured on the MacArthur-Bates Communicative Development Inventory) differed significantly between groups at followup.

DIR/Floortime-based interventions. The DIR/Floortime model focuses on understanding a child's challenges, strengths, and level of development to target relationship building, social and communication skills, and problem solving appropriate for a given child. DIR/Floortime programs typically include a variety of therapies including speech and language therapy, occupational therapy, and incidental and naturalistic teaching that follows a child's interests.

One fair quality RCT (Casenhiser 2013) evaluating an intervention based on the DIR/Floortime model and integrating parent training compared with community-based treatments reported no significant between group differences for language outcomes assessed using the Preschool Language Scale-4 or Comprehensive Assessment of Spoken Language.[128] Initiation of joint attention was significantly improved in the DIR/Floortime group compared with control at followup ($p<0.001$, effect size=1.02).

More than Words. The More than Words model emphasizes building parents' skills in interacting with the child and engaging the child in communication opportunities arising in activities or events occurring in the child's natural environment. The program focuses on teaches parental sensitivity to a child's developmental level and creating opportunities to promote communication. It targets two-way interaction, social communication skills, and improving understanding of language.

One RCT (Carter 2011) of fair quality compared language outcomes measured on both the Early Social Communication Scales and through coding of parent-child play for children receiving More than Words and children receiving treatment-as usual.[134] Language outcomes were not significantly different between groups at followup.

Pivotal Response Training (PRT). PRT focuses on "pivotal areas" of language and communication skill development including responsiveness, initiating communication, motivation to communicate, and self-regulation to attempt to produce wider improvements in language and other skills. PRT uses ABA techniques such as incidental teaching in natural environments, reinforcement, and motivation to promote communication.

One good quality RCT (Schreibman 2013) evaluated PRT compared with the PECS, which uses picture cards to promote communication in children with limited spoken language.[140] Both interventions targeted language use, either verbally or using picture cards. Language outcomes

measured on the Mullen Scales of Early Learning, Expressive One-Word Picture Vocabulary Test, VABS communication domain, and MacArthur Communicative Development Inventory did not differ significantly between groups, though both groups improved from baseline to followup. At followup 78 percent of children in both groups had acquired at least 10 spoken words.

Studies of Play-Based Interventions

Among the 12 studies addressing play-based approaches, we considered three studies to be of good quality[154-157] and ten of fair quality.[142-153] Studies were conducted in the United States[143,145,147-157] and Europe[144,146] and included a total of 483 participants between the ages of 21 and 82 months. Intervention duration ranged from 6 to 16 weeks; three studies reported long term (≥12 months post-intervention) followup of participants.[147,149-152,154] While all studies used approaches incorporating focused interactions directed by teachers or interventionists or parents/caregivers, studies typically addressed outcomes related to joint attention, play, imitation, or child/parent communication (**Table 9**). Participants in play/interaction studies often received other early intervention services in addition to the targeted intervention, making disentangling effects of the intervention difficult, but studies reported positive effects for preschool children with ASD, particularly when targeting joint attention skills themselves as well as related social communication and language skills.

Ten of 12 studies reported outcomes related to joint attention, and outcomes were significantly improved in the treatment group compared with the control group in nine; outcomes did not differ between groups in one study. Data are more limited about the ability of play-based studies to improve broad developmental skills (such as cognition, adaptive behavior, and ASD symptom severity) beyond communication and language gains over time. In one study, training regarding imitation skills showed positive results in improving not only imitation skills, but potentially other social communication skills such as joint attention as well. Studies focused on changing parental responses to child behavior and communication reported limited effects on language.

Key Question 4a. What Is the Effect of Intervention Timing (by Age and in Relation to the Establishment of a Definitive Diagnosis) on Treatment Outcomes?

In some early intervention studies, younger age has been associated with greater improvement. For example, greater language gains were seen in children who were younger with lower functioning levels at baseline in one RCT of an approach incorporating parent training.[138] In a retrospective cohort study of a community-based early intervention program, outcomes were related to age at enrollment, treatment duration, and higher baseline adaptive scores. A significant interaction emerged between age at enrollment and group membership, with younger starting age influencing outcomes for the treatment group but not the waitlist control.[109] Findings are not entirely consistent across age groups. Such differences may reflect variations in subgroups of participants responding differentially to treatment due to factors that are not fully understood or described in the research. Further as children develop, the specific targets of intervention often shift (e.g., studies at later emphasizing adaptive/social outcomes over

IQ/language targets) or samples participating in treat may reflect differential symptom profiles.

Additional analyses of some children in this earlier Canadian study[109] (overlap not clear) assessed the effects of baseline age and IQ on cognitive and adaptive outcomes in 207 children, and, in a separate analysis of matched older and younger children, effects of baseline age on the same outcomes.[141] In the initial retrospective analysis of 207 children, participant ages at intake ranged from 2 to 14.5 years, IQ from 10 to 104, and mental age from 3 months to roughly 7.5 years. Higher baseline IQ and younger age were significantly associated with greater cognitive rate (pre-post change in mental age/time in intervention) and with higher IQ at followup (all $p<0.001$), but change in IQ was not significantly associated with higher initial IQ. Higher baseline IQ was also associated with higher adaptive behavior scores at followup ($p<0.001$), but age was not a significant predictor. Longer duration of intervention was associated with slower rate of IQ and adaptive behavior development (p values ≤ 0.01); however, as this analysis was not prospective, the children who received more intervention could have been making slower progress. In the analysis of older (n=60, age 6 to 13.58 years at baseline) and younger (n=60, age 2.08 to 5.92 at baseline) children matched on developmental trajectory (i.e., number of intervention hours, baseline IQ and adaptive behavior), younger children had significantly better followup IQ outcomes compared with the older group. Younger children gained an average of roughly 17 IQ points (effect size=0.80) while older children gained an average of 2 points. Cognitive rate improved significantly for younger (effect size=3.19) but not older children. Both groups improved over time in adaptive behavior, but differences between groups were not significant (improvement of 4 points in younger children and 5 in older).

Key Question 4b. What Is the Effect of Severity of ASD (as Measured in Each Study) on Treatment Outcomes?

ASD Severity and Diagnoses

In some studies, children with lower symptom severity or less severe diagnoses improved more than participants with greater impairments. In an RCT assessing ABA-based early intervention, lower baseline ASD severity was associated with parent-reported cognitive and adaptive growth for children who received eclectic vs. ABA intervention, but not with standardized test scores.[105,106] A prospective cohort study of preschool-based early intensive intervention reported that children in the early intervention group with PDD-NOS or Asperger diagnoses (but not autism) had greater gains in overall adaptive behavior, communication, and daily living skills.[107]

Adaptive Behavior

Studies reported mixed findings related to outcomes associated with baseline adaptive behavior. In one retrospective cohort, positive outcomes in both the early intervention and the waitlist control groups were related to higher baseline adaptive scores.[109]

Language/Communication

The impact of language skills and attention to objects (vs. people) were assessed in two studies. In one RCT of the More than Words program, the treatment group showed differential effects on child communication depending on children's baseline object interest; children with lower levels of baseline object interest had greater growth in communication skills, whereas children with higher levels of object interest showed attenuated growth.[134] In another study of play-focused intervention, children with baseline expressive language abilities below 14.7 months showed greater gains in language in the intervention group vs. control (effect size=0.25 for 24 children with low language skills).[147]

Key Question 5. What Are the Harms of Treatment for ASD in Young Children?

No studies of behavioral interventions reported harms.

Chapter 4. Discussion

Summary of Evidence

This systematic review assessed the evidence for screening and treatment for autism spectrum disorder (ASD) applicable to primary care settings. We explicitly focused on studies examining screening strategies for young (<36 months of age), unselected populations (i.e., universal screening). We identified 17 unique screening studies in primary care populations consisting of approximately 13,884 participants who had not yet been identified as having a developmental concern. These studies are reported in 22 papers.[39,58,59,61-67,69-79,101] Although several ASD screening tools have been evaluated, only a small set of related tools have been adequately studied to provide data on young, unselected children (**Tables 2** and **3**). Most studies have evaluated versions of the same tool, the M-CHAT/F and M-CHAT-R/F. These screening approaches have been studied in more than 45,000 children, of which 76 percent (n=35,900/47,472) were in the United States and another 24 percent (n=11,572/47,472) used translated versions in Japan, Spain and Sweden. Other screeners have either already been demonstrated to provide unsatisfactory results or are not typically used in the United States, so the discussion below focuses on M-CHAT/F and M-CHAT R/F, which is the most likely screener to be used in the target population. The most current version is the M-CHAT R/F.

Findings for Well-Studied ASD Screening Instruments

Modified Checklist for Autism in Toddlers, Revised (M-CHAT-R) and Modified Checklist for Autism in Toddlers, Revised With Follow-Up (M-CHAT-R/F)

The M-CHAT-R and M-CHAT-R/F have been studied in more than 30,000 low-risk toddlers in heterogeneous community-based primary care settings. Positive predictive value ranges from 48 percent to 54 percent to accurately identify children with ASD between 16 and 30 months of age. In both versions of the screening approach, the proportion identified is lower than the recognized population prevalence, as calculated by the CDC. However, in one study that specifically assessed it, most children who screened positive had not yet been identified as being at risk of ASD by either their parent(s) or pediatrician. Specifically, as a result of screening, 60 percent of identified cases of ASD were found before parents expressed concern, 50 percent were identified before pediatrician concern, and in only 20 percent of cases were both the parent and pediatrician concerned before the screener identified a problem. These percentages are not mutually exclusive.

Both studies use a low risk population, excluding children already identified with developmental delay. It is unclear how the instrument operates in mixed populations that include children at increased risk for ASD. The study authors report that almost all children (cumulative 98%, including the autism diagnoses) were identified to have “actionable” concerns but, they do not provide detailed data other than general categories of diagnoses (e.g., language delay).

Of the two revised versions, the M-CHAT-R/F was most recently studied. It was developed to

addresses potential concerns about over-identification and misuse of the tool, specifically an over-reliance on the questionnaire without the follow up to validate results. Therefore, while the M-CHAT-R required a separate follow up, the M-CHAT-R/F packages the initial assessment and followup together as one two-step process.

None of the studies followed either the complete sample of screened negative children or a truly random sample in order to assess missed cases. Although true population estimates of how many children with ASD may screen negative on the M-CHAT or M-CHAT/F procedure are unavailable, comparing the proportions of children identified through the screeners with population estimates suggest this does occur. Given the prevalence rate of 1 in 68, as well as the resource-intensive nature of conducting the gold standard evaluations (i.e., in-person, time-intensive diagnostic assessments by skilled behavioral professionals sometimes lasting a full day or more) in screen negative children, almost no study has followed a large enough sample to truly assess false negatives on a population level.

The largest and most recent study followed up a quasi-random sample of 375 children who screened negative, of whom only 6 were eventually diagnosed with ASD (1.6%). An additional 42 out of 64 children who screened negative but were subsequently flagged by their pediatrician for further examination were also diagnosed with ASD, for a total of 11% (48/439) identified false negatives. The false negative rate for those who were identified as screen negative by the screening methodology and by physician observation was very low (1.6%). Although the methodology is imperfect, it does offer reassurance that the screening tools are unlikely to miss large numbers of children.

One additional approach to understanding potential under-identification, which has significant inherent methodological limits, comes from examining ASD identification rates relative to current prevalence estimates. The best study of the M-CHAT identified 1 in 208 children who participated in the initial screen while the M-CHAT-R/F identified 1 in 153. These numbers appear modest in comparison to the current prevalence rate reported by the CDC of 1 in 68; however, it is important to note that screening studies have been conducted within research programs, rather than standard clinical care. This means that families must opt-in and resources are expended to track, assess, and confirm concerns about ASD. Prevalence estimates applying simple corrections for attrition from these same studies range from 1 in 98 to 1 in 109, numbers that are closer to the current estimated prevalence rate (**Table 4**).

Given the pervasive and ongoing nature of developmental disorders, however, false negatives are primarily a concern among borderline cases, such as those that fail one portion of the screening process. Rather than follow up with all screen negatives, some studies followed up with borderline cases that failed a portion of the screening process (e.g. failed the screener but not the interview).

Other Tools

There are preliminary data suggesting that screening at very young ages (e.g. 12 months of age) with the First Year Inventory (FYI) might identify some children with social communication concerns. However, data have not yet been available to calculate estimates of psychometric

properties of this instrument in low-risk populations. The Infant Toddler Checklist (ITC) is designed as a general developmental screener for communication development that has also been used to screen for ASD. The ITC identifies many children with actionable developmental concerns, but using it in isolation may label ≥10 percent of the total low-risk population as at risk for ASD or other communication disorders. Identification of children via ITC screening may occur before parents and clinicians raise concerns.[39,63,76] Preliminary evidence suggests that children with ASD screened with the ITC may be referred for and receive higher levels of service than children with other developmental concerns.

There are some emerging data from non-U.S.-based programs and adaptations (Early Screening for Autistic Traits [ESAT], FYI-Lite, Social Attention and Communication Study [SACS], Young Autism and other developmental disorders Check-up Tool [YACHT]) regarding the potential for combined screening/surveillance and prospective strategies involving community health visitation and trained providers (e.g., visiting home nurses, behavioral health screening) to identify ASD at early ages. However, these tools have not yet been studied in U.S.-equivalent primary care settings and practices.

Screening in Special Populations

We did not identify any studies in primary care screening populations that specifically assessed the effect of individual risk factors such as prematurity or sibling status on performance characteristics of ASD screening tests. Our *a priori* review approach also precluded inclusion of screening studies in pre-selected populations, such as siblings of individuals with ASD. Later-born siblings of children with ASD are estimated to see an incidence rate between 7 and 19 percent. Screening siblings may be particularly challenging using measures that rely on parent report because these families evaluate their second child's behavior and development through a different lens. Further, these subgroups may have very complex profiles of developmental skills.

Within the NICU graduate population, where ASD is estimated to occur 5 to 8 times more often than in the general population,[158] it is difficult to discriminate children at-risk of ASD from children at risk of other neurodevelopmental conditions. Several studies conducted to date have reported positive ASD screening rates of 10 to 25 percent when the M-CHAT is used in toddlers who were low birth weight babies (see Stephens et al, 2012;[159] Limperopoulos et al, 2008;[160] Kuban et al, 2009[161]).

Screening by Age

No studies focusing exclusively on primary care screening systematically examined the performance characteristics of early screening tools for children at different ages within our time frame (0-36 months).

One study,[70] representing a subsample of high- and low-risk children overlapping with populations reported in other studies in this review,[101] attempted to examine screening characteristics of the M-CHAT at different ages. PPV was higher for children above 24 months. However, many children were excluded from this sample as they had already received an ASD

diagnosis. As such, the sample may not reflect true population estimates of PPV by age in pediatric settings.

Indirect evidence suggests that studies in which the target age was entirely or primarily below 16 months (ESAT, ITC, FYI) report both under-identification (low sensitivities and PPV) and over-identification (low specificity). Some instruments (e.g., M-CHAT) have not yet been thoroughly studied in very young populations so performance characteristics are unavailable. Some instruments also recommend different cutoffs based on age (e.g., ITC) whereas some screeners rely on the same algorithm to determine risk across a wide span of development. In both cases there may be differences in performance characteristics that vary by age, but data are not yet available to assess these differences.

Outcomes of Screening

Children who were screened have had earlier ages of concern, referral, and diagnosis compared with a) parent and clinician ratings and b) population estimates of age of diagnosis.[39] Only one study to date has examined age of identification and service entry with a randomized controlled study of developmental screening in comparison to surveillance,[162] but this study did not provide performance characteristics for an ASD diagnosis specifically, so was excluded from our analysis. It did, however, use the M-CHAT in addition to general developmental screening, with screening associated with earlier age of identification with developmental delay, earlier referral for service, and increased service access through early intervention systems. Children in the screening group had a 59% to 68% shorter time to identification of their particular diagnosis, and a 64% to 70% shorter time to referral for services. We did not find a corollary study for ASD alone.

In terms of access to services, children screened at risk and later diagnosed with ASD receive higher levels of service compared to children without such diagnosis. However, there have not been any systematic prospective studies that examine service access and the impact of service delivery for children screened at-risk after they are diagnosed. Some information suggests that children screening positive for ASD are referred for and receive higher levels of early intervention service than children identified via clinical surveillance.[39]

Outcomes of Treatment

Early Intensive Behavioral and Developmental Interventions

Evidence for the effects of early intensive behavioral intervention is primarily available for the outcomes of cognitive improvement and language (receptive and expressive), and the primary intervention is early intensive developmental and behavioral intervention. Twenty of the 30 studies of early intensive intervention assessed cognitive outcomes. Eleven of these 20 studies demonstrated significantly greater effects on cognitive outcomes among children in the treatment group compared to children who received eclectic community-based care (the comparator in 13/20 studies assessing cognitive outcomes) or other comparators. Among 15 of 17 studies of direct provision to children reporting cognitive outcomes, 10 studies demonstrated greater

changes in cognitive outcomes in the treatment group relative to the comparison arm. Among five of 13 parent-focused studies reporting cognitive outcomes, one reported significantly improved outcomes in the treatment group compared with control.

For example, in the good quality RCT of the Early Start Denver Model, cognitive outcomes for children in the ESDM group improved a mean of 17.6 points from baseline compared with a 7 point gain in the control group ($p=0.044$) after 2 years. Compared with children in preschools with teachers who received LEAP manuals and no specific instruction, children in preschools with teachers who received LEAP coaching and training gained significantly on cognitive measures (mean gain of 8.9 points vs. loss of 1.8 points on the Mullen Scales of Early Learning Early Learning Composite, $p<0.01$). In studies of the UCLA/Lovaas-based approach, children in the treatment group were significantly more likely to achieve cognitive scores in the average range at followup than the eclectic group (7/13 children vs. 2/12, $p<0.05$).

Because of the variability in test item content (e.g., nonverbal problem solving, verbal reasoning, motor skills) and scoring protocols (e.g., different cut-off points for basal and ceiling scores), it is difficult to directly translate numerical changes on a cognitive or language scale to specific developmental changes within an individual child. However, these overall scores, particularly on cognitive scales, can make a difference regarding a child's educational or diagnostic classification, particularly with regard to a co-occurring diagnosis such as an Intellectual Disability. A boost of a few points in a cognitive score can make the difference between a Mild Intellectual Disability versus Borderline Intellectual Functioning versus a score in the Average range. The subsequent impact of these classifications on environmental or intervention factors, such as classroom placement, could offer children additional opportunities over time that, again, are difficult to quantify but are nevertheless important.

Another possibility is that shifts in cognitive scores partly reflect a child's increased ability to attend to an adult, understand spoken instructions, sit at a table, and respond in the correct way to a spoken question—all important skills that can help children participate in, and learn from, structured and unstructured social situations. Many young children with ASD have difficulty following the format of a standardized assessment in ways that reflect their underlying social-communication and play vulnerabilities. For example, instead of following the examiner's instruction to match objects, they may put the objects in a line, spin them in their peripheral vision, or not engage with the objects at all, instead exploring other sensory properties of the room. Increased cognitive scores, therefore, likely indicate to some degree a child's increased ability to attend to social interactions, one of the key deficits in ASD.

Given that these tests reference skills and mark change in relation to normative samples of typically developing children substantial variation in the population of interest can often appear more moderated (i.e., children gaining two years in developmental age and children gaining only a month in developmental age) and or profoundly effects of developmental floors of instruments. Nevertheless they do reference change regarding normative development. As such, changes relative to standard scores as such indicate not just developmental progression, but progression outpacing what would be expected during typical developmental progress (i.e., more than a year of developmental progress within a chronological year). In this regard, there can be potential

value in describing the normative skill gains in corresponding to developmental age equivalents in order to better understand the clinical functional impact, or lack thereof, in interventions. Twenty-eight of 30 studies assessed language outcomes. In 13 of these 28 studies, outcomes were significantly improved in the treatment arms vs. the comparison arms. In 19 of these 28 studies, the comparator was eclectic treatment, and outcomes were significantly improved in the treatment vs. comparison group in 10 of these 19.

Data on language outcomes was more mixed and as follows. ESDM was associated with statistically significant increases in language acquisition relative to community eclectic treatment. Two studies of LEAP assessed language outcomes with mixed results (improved outcomes in the full LEAP training arm compared with control in one and no significant differences among groups in another). Among the studies of UCLA/Lovaas-based approaches, two studies demonstrated significantly greater improvements on the Vineland Adaptive Behavior Scales Communication domain with treatment versus eclectic controls; results were mixed (one study positive and one neutral) when outcomes were measured on the Reynell Developmental Language Scale, and one study reported that UCLA/Lovaas-based treatment was associated with improvements in receptive but not expressive language, again when compared to eclectic comparators.

Clinical Implications of Changes in Cognitive and Language Measures

In order to benchmark scores we provide descriptions of the age equivalent and corresponding item sets and abilities from the Mullen Scales of Early Learning in the Dawson (2010) study of ESDM and eclectic, community-based intervention. As the study was an RCT of two active treatments (e.g., manualized intensive developmental/behavioral intervention vs. eclectic community treatment) we can present a description of the indexed change for both samples.

Cognitive Outcomes

In the ESDM group at baseline, children at an average of age of 23.9 months had an Early Learning Composite (standard score [SS] mean=100±15) of 59.4 (1^{st} percentile). This corresponded to domain scores (T scores Mean=50±10) of 21.1 (1^{st} percentile) regarding language understanding (Receptive Language); 24.5 regarding language use (1^{st} percentile; Expressive Language); 33.2 (4^{th} percentile) regarding nonverbal problem solving (Visual Reception), and 33.9 (4^{th} percentile) regarding fine motor skills (Fine Motor).

Such baseline performance corresponds to receptive and expressive language skills equivalent to a 13 month old. Children in this span of functioning are likely to be able to voluntarily babble, potentially produce a very small number of single words (<8), but not yet successfully label objects, pictures, or use phrases. In terms of language understanding children at this level would be likely to understand simple contextualized instructions (e.g., 'no,' 'let's go'); however would not be able to follow simplistic novel directions (e.g., show me your nose/eyes, point to the cat, etc.). Nonverbal skills would correspond to 17 month old developmental level (Visual Reception) with Fine Motor skills at an 18 month old level. Children at this level of functioning would be able to explore objects with their hands, demonstrate object permanence, take small objects in and out of containers, and potentially place small inset puzzle shapes (e.g., circle,

square inset). Children at this developmental level would have challenges stacking and building blocks in imitation of a town or train, stringing beads, as well as sorting or matching simple shapes.

At 52.4 months the ESDM group had an average Early Learning Composite (SS Mean=100±15) of 78.6 (7^{th} to 8^{th} percentile). This corresponded to domain scores (T scores Mean=50, SD=10) of 40.1 (16^{th} percentile) for language understanding (Receptive Language); 38.6 for language use (12^{th} percentile; Expressive Language); 41 (18^{th} percentile) for nonverbal problem solving (Visual Reception), and 33.5 (4^{th} percentile) for fine motor skills (Fine Motor). Children in the other treatment arm (community treatment) did not exhibit statistically significant difference on their cognitive profile.

The increase in skills seen in the ESDM group represents 30 to 31 months of progress in language skills, 29 months in nonverbal skills, and 24 months of progress in fine motor skills. Children with developmental language skills on this level are commonly able to use three to word sentences, repeat number sequenced, count, answer simple questions, follow novel two step-unrelated commands, identify colors, shapes and sizes.

In the community treatment group at outcome, children had new developmental age equivalents of approximately 37 months, Receptive Language; 36 months, Expressive Language; 41 months, Visual Reception; and 40 months, Fine Motor. This represents 24 to 25 months of progress in language skills, 25 months in nonverbal skills, and 23 months of progress in fine motor skills.

Simply charting the corresponding age equivalents to the 17.6 point differential at outcome in the treatment arm (i.e., comparing age equivalents of baseline cognitive scores and outcome cognitive scores) this amounts to approximately 5 to 8 months of increased developmental acceleration during a 2 year span. It is important to note that this change is in addition to typical developmental progress of a normative sample leaving open the possibility that this change may far exceed the developmental progress of untreated ASD groups who may actually progress at a much slower rate in terms of cognitive skills. Thus, both groups made progress, but the active treatment group (ESDM) made significantly greater gains, as described above.

These are average effects, and the range of effects is wide. Not all experience such positive outcomes. Nonetheless, studies of long term outcomes for individuals with autism up to 19 years of age point to early cognitive scores as a primary predictor of success, including maximizing independence, so research continues to focus on early interventions that can move the bar on cognitive scores. New research suggests that not only cognitive scores themselves, but shifts in cognitive scores at the earliest ages, is associated with positive outcomes as well.[163]

As another example, a 25-point gain in cognitive scores from a baseline of 69 (which is within the range of Mild Intellectual Disability) would yield an IQ score of 94, within the Average range. This reflects an improvement of almost two standard deviations within most standardized cognitive instruments, shifting a child's score to a significantly higher percentile rank relative to other children who take that test. This score increase may reflect improved verbal and nonverbal reasoning skills (e.g., defining words or relationships between words), visual processing skills (e.g., matching pictures, replicating block designs), working memory (e.g., remember chains of

letters and numbers), or processing speed (e.g., identifying a certain number of targets within a given time period), to name a few of the many areas that cognitive tests assess. It may also reflect an improved ability to respond to standardized items and questions, with the potential for broader understanding of social rules and ability to learn and implement adaptive behavior tasks. For example, a child with the ability to point to pictures as part of a cognitive test may also be able to point to pictures as part of a communication system. A child who can sort objects may respond to structured work systems to promote his or her independence across environments. A child who can verbally define a word also understands the meaning of that question which, in turn, could translate into understanding adult instructions and verbalizing thoughts and needs. All of these skills in turn promote participation in classrooms, therapy sessions and, in the long term, employment settings. Although there is certainly not a one-to-one correspondence between cognitive scores and these types of clinical outcomes, increases in scores may reflect increased ability to understand, implement, and generalize these skills.

Language Outcomes

Regarding language measures, increases in scores may translate more directly to everyday applications. Increases in expressive language (ability to use words and phrases; may also include gestures) and receptive language (ability to understand words; e.g., follow instructions, find named objects) can have concrete meaning for families, even if the increases are small. For example, a child without any words who does not point may scream or use a parent's hand as a tool to request help. If that same child is able to point to a picture of what he wants, or use a handful of single words (e.g., juice, Mama), it can reduce the child's frustration, thereby promoting social engagement. The acquisition of verbal skills may also have a strong psychological impact on parents, many of whom identify a speech delay as their first concern about their child's development. Similarly to IQ, language skills also have impacts on children's classroom placements and functioning, including the kinds of communicative supports necessary for them to participate (e.g., Alternative and Augmentative Communication Devices, PECS) and understand what is going on, including daily schedules, classroom rules, and social interactions with peers.

Play-Based Studies

Studies incorporating play or interaction-based elements typically targeted joint attention skills and included younger children (21 to 82 months). Joint attention skills are thought to be pivotal skills central to the etiology and neurodevelopmental sequelae of ASD.[164,165] Pivotal skills, or social communication building blocks, are intended to enable children to learn from their natural social environment so that development continues as a result of having acquired the skill, even in the absence of continued treatment.[166] At a basic level, "joint attention" refers to specific skills that involve sharing attention, particularly visual attention, with others (e.g., pointing, showing objects, coordinating gaze, and responding to such bids). These exchanges enable young children to socially coordinate their attention with other people to more effectively learn from their environments. Fundamental differences in early joint attention skills, which emerge and develop within and across infancy, have been linked to the deleterious neurodevelopmental cascade of ASD, and successful treatment of these deficits could substantially improve other skills across settings and the lifespan.[151,154,165]

Participants in play/interaction studies often received other early intervention services in addition to the targeted intervention, making disentangling effects of the intervention difficult, but studies reported positive effects for preschool children with ASD, particularly when targeting joint attention skills themselves as well as related social communication and language skills. Data are more limited about the ability of play-based studies to improve broad developmental skills (such as cognition, adaptive behavior, and ASD symptom severity) beyond communication and language gains over time. In one study, training regarding imitation skills showed positive results in improving not only imitation skills, but potentially other social communication skills such as joint attention as well. Studies focuses on changing parental responses to child behavior and communication reported limited effects on language.

Moderators of Outcomes

Existing studies do not predict which children are most likely to benefit although effects have a wide range, with some children exhibiting substantial improvement and others not at all. One of the challenges in this literature base is that many studies do not adequately characterize the participants; nor do they provide subgroup analyses that might be helpful in describing which children benefit most. However, higher baseline cognitive scores were associated with less improvement in some interventions. Since interventions are often targeted to individual baseline deficits, global measures may not capture individual effects. And, even children who demonstrate clinically significant improvement in response to intervention often continue to display substantial impairment across some areas of functioning.

Indirect evidence suggests that earlier intervention may result in better outcomes. In ABA-based and parent training studies, younger age at enrollment emerged as a predictor of improvement in cognitive skills and language comprehension and production.[126,129,138] In early intervention studies, younger age was associated with greater improvements: greater language gains were seen in children who were younger with lower functioning levels at baseline in one RCT of an approach incorporating parent training.[138] Another study assessing parent-delivered ESDM reported greater increases in developmental quotient scores in children under 24 months of age in analyses combining the ESDM group with the control group, which received community-based treatment (effect size=-1.20, p=0.002).[129]

In a retrospective cohort study of a community-based early intervention program, outcomes were related to age at enrollment, treatment duration, and higher baseline adaptive scores. A significant interaction emerged between age at enrollment and group membership, with younger starting age influencing outcomes for the treatment group but not the waitlist control.[109] However, these findings are not entirely consistent, and one study comparing preschool-delivered intensive early intervention and treatment as usual reported larger adaptive behavior gains for older children in the early intervention group.[107] Such differences may be related to subgroups of participants who respond differentially to treatment due to factors that are not fully understood or to developmental issues limiting specific interventions, such as lesser ability to develop adaptive behaviors before a certain age; further, age at diagnosis may also be associated with severity, which could also confound response to treatment.

Challenges in the Evidence Base

A remaining significant challenge to interpreting the early intensive intervention literature relates to how interventions are described and implemented. Although researchers are increasingly attempting to manualize (develop standardized treatment manuals) approaches as well as operationalize and measure treatment fidelity, most of the body of literature categorized in this report as "early intensive behavioral and developmental intervention" remains an eclectic grouping. This category of interpretation presently groups different treatment approaches (i.e., developmental, intensive behavioral, center based, and combinations), intensity (12 hours over 3 months vs. 30 hours over 1 week), and duration (weeks to years); varied inclusion and baseline assessment criteria; children of varying ages (intake age ranging from 18 months to 7 years); and many different outcome measurements over different periods of time (weeks to years). There are potent intrinsic challenges to manualizing intensive interventions to be delivered over months and years for a heterogeneous patient population. However, recent progress toward this end has shown that children will often respond differentially to early intensive approaches.

Few studies directly compared the effects of well-controlled treatment approaches, instead comparing interventions to non-specific "treatment as usual." Additionally, little data on the practical effectiveness or feasibility of these treatments beyond research studies exist, and questions remain about whether reported findings would generalize on a larger scale within communities. Similarly, no studies in this category reported harms of intervention in terms of child, family, or system impact.

Findings in Context of Other Recent Reviews of Behavioral Interventions

We rated three meta-analyses evaluating early intervention for children with ASD as good quality;[167-169] we also summarize two overview meta-analyses (not quality rated) addressing early intervention.[170,171] Of the 30 early intensive behavioral and developmental studies described in our review, four[117-119,122,123] were included in three of the prior reviews[167-169,171] summarized here, and two[19,111] were included in four[167-169,171] of the reviews. One additional study[107] was included in one prior review.[169] Although none of the prior reviews used exactly the same set of studies, and most used many fewer studies than ours, these reviews were in general agreement with our assessment of the literature, finding significant positive effects of early intervention on cognitive and language outcomes, and also noting that the quality of the literature was not optimal and should be improved in future research.

One Cochrane review compared early intervention with treatment as usual and included RCTs or controlled trials with participants under 6 years of age at intake.[167] The review included five studies (one RCT) with a total of 203 participants (mean age range: 30.2 to 42.5 months). The investigators rated all studies as having high risk of bias (low overall quality) and found positive effects for early intervention on all outcomes. Mean difference effect sizes were 0.76 for cognitive scores (95% CI=0.40 to 1.11, $p<0.0001$), 0.69 for adaptive behavior (95% CI=0.38 to 1.01, $p<0.0001$) and ranged from 0.42 to 0.74 for measures of communication, socialization, and daily living skills (p values=0.0005 to 0.03). Tests of heterogeneity and small sample sizes

precluded assessment of moderators of effects.

One meta-analysis of ABA-based interventions included studies with at least five children with ASD receiving at least 10 hours of intervention per week for 45 weeks.[169] Twenty-two studies met criteria and assessed outcomes including IQ, receptive and expressive language, and adaptive behavior (Vineland Adaptive Behavior Scales composite and domain scores). Studies included 323 patients (mean age 22.6 to 66.3 months, 55.6 to 97% male). Study quality was low to moderate, ranging from 1.2 to 3.6 on a five point scale (mean 2.5). Thirteen studies had control groups (six with random/quasi-random assignment). Positive effects were associated with ABA-based intervention in 18 studies assessing the outcome with a pooled effect size of 1.19 (95% CI: 0.91 to 1.47, $p<0.001$). Similarly, ABA was associated with positive effects on language (general, expressive, and receptive, effect sizes from 1.07 to 1.48) and adaptive behavior (communication, socialization, motor skills, daily living skills domains as well as composite scores; pooled effect sizes ranging from 0.61 to 1.45). The effect size for the composite score was 1.09 (95% CI: 0.70 to 1.47, $p<0.001$), and total treatment duration was associated with better adaptive behavior and language outcomes but not IQ. Results restricted to studies with control groups were consistent with results for all studies across outcomes. Across outcomes, effect sizes were generally slightly better for clinic-based approaches vs. parent-delivered. The investigators note the potential for publication bias for the cognitive and language outcomes and the adaptive behavior domains of communication and socialization.

Another meta-analysis of ABA-based early intervention included 11 small comparative studies (one RCT) with 344 children with ASD (mean age 33.56 to 65.68 months, 65.7% male).[168] The mean quality of studies as rated on the Downs and Black scale was 24.65 out of 32 (range 23-27). The early intervention group had greater gains on all variables assessed compared with control group participants, with full scale IQ improving by 11.98 points over improvements in the control group. Receptive and expressive language scores for the early intervention group compared with control each improved by more than 13 points, while improvements on Vineland Adaptive Behavior Scales domain scores ranged from 4.96 to 10.44 points. Total effect sizes for daily living skills improvements were moderate (0.68) and were large for improvements in IQ, language, and adaptive behavior (effect sizes ranging from 0.91 to 2.00). The authors noted some evidence of publication bias. **Table 15** outlines key characteristics of these early intervention meta-analyses.

A sequential or cumulative meta-analysis compiled data from 15 studies rated as adequate or high quality in five previously published meta-analyses (Eldevik 2009, Makrygianni 2010, Peters-Scheffer 2011, Reichow 2009, Spreckley 2009).[171] The 15 studies included 263 children with ASD. The sequential meta-analysis found a medium treatment benefit for early intervention vs. comparison interventions for the outcomes of intellectual functioning, language, and adaptive behavior. The magnitude of treatment benefit varied for outcomes when assessing pre- to post-differences in the early intervention group. For IQ, the standardized mean difference effect size for group differences was 0.61 ($p<0.001$) and the pre to post differences in the early intervention group was 0.71 ($p<0.01$). Between group effect sizes for adaptive behavior and language were also considered medium (0.60 and 0.72, respectively, p values <0.001). Pre- to post- effect sizes were for adaptive behavior (0.35, p=NS) and language (0.69, $p<0.05$) did not reach sufficiency and could not be considered as providing evidence of medium pre to post treatment benefit. The

authors note that meta-analyses for pre to post differences in adaptive behavior and language were underpowered.

An overview of four of the same meta-analyses noted above plus one additional (Virues-Ortega 2010) described methodological limitations across the meta-analyses.[170] Limitations included small sample sizes in included studies, inclusion of nonrandomized studies, lack of standardized control groups, errors in interpretation of studies, and variations in the early intervention approaches assessed. Four of the five meta-analyses concluded that early intervention was an effective approach. For IQ, the weighted mean effect size across meta-analyses ranged from 0.38 to 1.19 and from 0.30 to 1.09 for adaptive behavior. Despite the need for additional research, particularly in understanding effective treatment components and child characteristics associated with optimal outcomes, the authors conclude that early intervention can produce significant effects on cognitive skills and adaptive behavior for many young children with ASD.

Table 16 summarizes studies identified and findings across key questions addressed in the review.

Review of Contextual Issues

Contextual Question 1. Are Screening Resources Currently Adequate to Support Routine Screening in Primary Care or Affiliated Care Systems?

Current estimates suggest that approximately 42 to 55 percent of pediatricians regularly screen for ASD in toddlers,[2], but providers are much less likely to screen toddlers from underrepresented ethnic and language groups (10% to 29%).[3] While there is substantial variability in utilization of tools, limited information is available regarding characteristics of providers who are more likely to screen for ASD.

Availability and Training in Tools

The M-CHAT and companion interview, M-CHAT/R-F, as well as the ITC are available for free download and use in primary care settings. The original M-CHAT and ITC are available in numerous different languages (M-CHAT: 30 languages other than English at the time of this report). Behavioral coding, assessment, and interview procedures used to follow-up ITC screens have varied substantially, are not always clearly stated within protocols,[63] and may require additional resources and training to accomplish. Administration and scoring of the M-CHAT, M-CHAT/R-F, and ITC instruments requires no specific training and scoring takes less than 5 minutes. Utilization of the follow-up interviews embedded within M-CHAT and M-CHAT-R/F and specific referrals for children flagged with concerns may take substantially longer. However, this interviewing is designed to be employed only when concerns are in a questionable range (i.e., 3 to 6 failures; automatic failure if self-report score >7) which somewhat limits the time needed (estimated at 5 to 10 minutes by authors) to accomplish the interview.[61,101] Although designed to be used by extenders with limited resources, there may be some training and capacity issues related to how initial results of screeners and next steps would be conveyed to caregivers.

The SACS, YACHT, and ESAT all use assessment procedures that are not freely available. Their training requirements for use outside of specific research protocols are unclear (i.e., trained nurses and home visiting psychologists).

Practice Supports for Screening

The "Autism Case Training (ACT): A Developmental Behavioral Pediatrics Curriculum" is available from the Centers for Disease Control and Prevention (CDC) and offers continuing education credit to providers accomplishing training. This series was written by developmental-behavioral faculty and fellows from 10 Maternal and Child Health Bureau Developmental-Behavioral Pediatrics Fellowship Training Programs and the CDC. In addition, the AAP has developed an ASD Resource Toolkit for care providers that supports the medical home in identification and care for children with ASD. Finally, Autism Speaks maintains resources for both practitioners and families designed to ensure ready access to information about ASD care.

Referral Resources

Children who screen positive as a result of ASD screening practices are often referred for further evaluation and intervention services, in addition to additional surveillance and care facilitated by the primary care provider. For many primary care providers, the most common referral pathway is a referral to Early Intervention (Part C) programs within their states.[45]

The Program for Infants and Toddler with Disabilities (Part C of Individuals with Disabilities Education Act [IDEA]) is a federal grant established by Congress. It assists states in operating a statewide program of early intervention services covering children ages birth through 2 years, 11 months of age. For a state to participate in the program it must assure that early intervention will be available to every eligible child and its family. Currently, all states and eligible territories are participating in the Part C program.

The current IDEA Statute (P.L. 108-446) contains many requirements that states have to meet. However, states have some discretion in setting the criteria for child eligibility, including whether or not to serve at risk children. As a result, service eligibility definitions differ significantly from state to state. Given this variability, not all children screening at risk for ASD are automatically eligible for services. However, all referred children are eligible for assessment through this program within a legally mandated time period. Eligibility does not hinge on a diagnosis of ASD and services are to be structured based on family goals. Available appropriate services for children eligible for Part C often differ dramatically by state, geographical regions within states, as well as by family and child. Historically, children with a specific ASD diagnosis have obtained higher intensity specialized intervention services than children without a diagnosis.[39,40] However, state Part C systems increasingly are incorporating specific ASD screens into their own eligibility process. They are also incorporating explicit mechanisms to follow-up on provider screens in order to provide eligibility for intensive levels of ASD intervention that may, in some cases, be initiated prior to diagnosis.[172] Many Part C systems also incorporate mechanisms for obtaining an ASD diagnosis into their eligibility and service process. Despite the variability across states, the age focus of IDEA on children up to 36 months is consistent with a focus on early childhood screening as addressed in this report.

Providers may refer screen positive children to private service providers (speech/language pathologists, occupational therapists, behavioral providers, center and school programs) in addition to, or in conjunction with, referrals to Part C systems. The availability and accessibility of such additional referral resources varies across communities and is difficult to catalogue.

Contextual Question 2. Are Diagnostic and Treatment Resources Currently Adequate to Provide Services to Children Who Screen Positive for ASD?

Concerns about lengthy waiting lists and/or over-referral of children who meet screening criteria for reasons other than ASD have been postulated as contributing factors relating to incomplete incorporation of ASD screening into pediatric practice[51,52] and ultimately to the delays in diagnosis seen on a population level, with many children diagnosed after age 4 in spite of accurate diagnosis being feasible at younger ages.[53,173] Currently, there is substantial variability across communities in terms of access to diagnostic and treatment resources.

A formal ASD diagnosis is not necessarily a specific requirement for the initiation of services through many programs (e.g., state- and federally-funded early intervention (Part C) and school (Part B) systems, adjunctive therapy services). However, access to subsidized early intensive behavioral and developmental intervention (i.e., supported by health insurance, school systems, early intervention systems) is often limited without a confirmed diagnosis, and children commonly receive lower levels of service in the absence of such a diagnosis. Although some Part C systems offer pathways for children to receive diagnostic clarification and subsequent services, disparities still exist in access to diagnostic assessments. For some children, the time from screening to diagnosis and treatment may be estimated in weeks; however, for other children waits for diagnostic assessment may be in excess of 6 months.[49,174]

Data from the most recent CDC prevalence study indicate that more children were formally diagnosed with ASD in the 2014 Autism and Developmental Disabilities Monitoring (ADDM) report than at any previous time point. At the same time, a significant proportion of children with ASD did not have a clinical diagnosis either in their medical or educational record, which suggests capacity is not currently meeting the need. Additionally, only 44% of children with ASD diagnoses via ADDM monitoring had received evaluations for developmental concerns prior to age 3.

There are limited data available on the numbers of families able to access services on a community level. Currently, all states and eligible territories participate in the Part C early intervention program for children ages birth to 3 years. This system presumably allows children to receive services based on risk prior to diagnosis as well as post-diagnosis, but these services may range in intensity and focus. Children who are over age 3 may have access to additional services through their school district, but the nature of services provided within these systems is variable.

Most U.S. states (estimated at 34)[175] have enacted ASD insurance reform legislation that provides for specific access to evidence-based intervention services through private insurance. The availability and accessibility of resources for referral varies across communities.

Contextual Question 3. To What Degree Does Evidence Indicate That Children Can Be Accurately Diagnosed With ASD, Using the ADOS and/or Expert Clinical Opinion? What Does the Current Scientific Literature Demonstrate Regarding the Persistence of an ASD Diagnosis?

ASD is a behaviorally based diagnosis with clinical diagnosis the gold standard in diagnostic accuracy; this diagnosis incorporates information from multiple sources and depends on the professional training and expertise of the evaluator. Information about medical and developmental history, cognitive and developmental functioning, and behavioral observation are crucial components of ASD evaluation. Complex cases are particularly reliant on information from a skilled clinician and caregiver.

Increasingly clinicians and researchers have specific tools to aid in assessing the core features of the disorder. The most well validated and commonly used measure in both research and clinical settings for diagnosing ASD is the Autism Diagnostic Observation Schedule (ADOS). This measure is a semi-structured assessment that rates key social communication skills and atypical behaviors associated with ASD. It has historically been considered crucial in research and is becoming more commonly used by clinical and educational professionals as well. The instrument is available for purchase from a U.S. publisher along with training materials and videos. The newest iteration of the ADOS, the ADOS-2, has sensitivity and specificity for diagnosing ASD from 91 percent and 94 percent for toddlers with few to no words and 88 percent to 94 percent for toddlers with some words.[176] Similar discrimination and sensitivity for children in the second year of life have also been documented in samples not affiliated with the authors of the instrument.[177] The administration of the ADOS-2 is one component of a full diagnostic evaluation for ASD. While it is considered the most common instrument in diagnostic assessment practice, it is not a necessary component for such a diagnosis. Information about medical and developmental history, cognitive and developmental functioning, and additional behavioral observations can be obtained via numerous methodologies and reports.

A growing body of work suggests that diagnoses of ASD delivered by an expert clinician using standardized assessment tools in predicting continued ASD diagnostic classification up to school age.[93,178-180] However, given some clinical recommendations for screening during the second year of life, there has been a growing need to understand how well we can diagnose young children and whether diagnoses in this younger age group are stable. Although most of the initial diagnostic accuracy evidence in children under age 2 was from studies conducted with experienced and highly trained staff in research settings,[181,182] there is some evidence that the ADOS-2 can be used for young children in community-based clinic environments.[183] Pediatric providers can be trained to accurately diagnose ASD[49,174] and clinicians using community screening protocols may be able to accurately identify ASD between 15 and 23 months of age for most children.[177]

Contextual Questions 4 and 5. To What Extent Does Over-Diagnosis and Under-Diagnosis Exist in Children Referred to Diagnostic Evaluation After Screening?

Our review found the CDC's ADDM Network provides some evidence of the degree to which there is over or under identification of ASD in children. In the most recent CDC review of health and special education records from 2010, roughly 20 percent of 8-year-old children who had been identified as having ASD through the ADDM classification process did not have a formal educational classification or clinical diagnosis in their medical or educational charts.[1] In some catchment areas, this number was as high as 43 percent. Many of these children were identified under a different classification than ASD for obtaining educational services. This suggests that on a population level, the collective medical and educational systems may not be completely or accurately identifying ASD at early ages.

ASD has recently been redefined in DSMV to include individuals across the spectrum, and understanding of the attendant variations in natural history is incomplete. Although historically ASD has been thought of as a lifelong condition, there is increasing awareness that some children at later ages appear to have more "optimal outcomes."[184] Some children identified with ASD at early ages may demonstrate improvements in functioning, both in relation to maturation as well as intervention, such that a diagnosis of ASD may no longer be appropriate. This idea of "best outcome" and/or "recovery" is not novel.[185,186]

Recent work suggests that even after receiving thorough, accurate diagnostic classification at young ages, there will be some individuals who no longer meet criteria for ASD and whose communication and socialization skills are on par with typically developing individuals, but research in this area is so new that estimates of how many individuals this includes are unavailable.[184,187] Nonetheless, preliminary numbers vary across studies. Although early estimates of "optimal outcomes" for children with ASD by Lovaas[185] reached 47 percent with intensive levels of treatment (40 hours per week of individual ABA for at least 2 years), more recent estimates of such outcomes are smaller and variable,[188] with some evidence that even in children with initially average intelligence, significant challenging behaviors across the lifespan may reduce ability to even complete standardized assessments by adulthood.[189] Smith et al. found that compared to a parent training group, children who received 30 hours of individual ABA per week (again, for a period of years) achieved significantly more regular education classroom placements (both with and without supports) than controls, but only 2/15 children in this group achieved regular education placement without support as well as IQs in the average range.[120] It is unknown whether such groups represent diagnostic errors, correct diagnoses in children whose developmental pattern encompasses significant improvements in ASD-related impairments, or are the results of accurate early diagnosis and treatment. A review of the adult outcome literature suggests that improvements like these have not been common,[190] but may become more common as cohorts of children receiving interventions are tracked to early adulthood.[163]

Contextual Question 6. Do the Outcomes of ASD Screening and Efficacy and Harms of ASD Interventions for Young Children Differ by Pertinent Subgroups, Such as Racial/Ethnic Minority, Low-Income, and Uninsured Children?

Subgroup data are largely unavailable at this time. Most of the highest quality intervention studies have not adequately included families from traditionally underserved backgrounds. Children from underserved communities are less likely than children from other communities to be screened,[3] in spite of data indicating that accurate primary care screening and referral are possible in diverse and traditionally underserved populations.[39,63,162] The effects of socioeconomic characteristics on treatment outcomes have not been well explored. However, some data suggest that families with lower SES may be less likely to access higher intensity interventions.[120] Families with greater annual incomes are more likely to have children diagnosed at young ages,[191] whereas children of parents with lower socioeconomic or educational status are less likely to be diagnosed promptly.[37,190,192-194] Research into the effects of racial and ethnic backgrounds on ASD diagnosis and intervention is similarly lacking. Existing evidence indicates that compared with children from other ethnic and racial groups, African-American children are more likely to receive an inaccurate diagnosis before being identified correctly as having ASD.[193] Additionally, CDC researchers have documented a historical trend for later identification of ASD in other racial/ethnic minority groups as well.[1]

Limitations of the Review

There are several significant methodological issues in the currently available screening literature. First, many studies of purported ASD screeners have actually been conducted in settings with limited relevance to primary care and potential universal screening practices. These clinical and convenience samples do not adequately demonstrate the psychometric properties of screeners in practice for several reasons: they do not usually include children with a range of other concerns, they rely on parent-report of symptoms in children in whom ASD has been identified, or they include only the subset of individuals presenting to these defined clinical and research settings rather than the entire referral sample. We excluded these studies from our review, but were therefore left with a small number of studies to include.

A second significant concern is the lack of data on, and adequate methods for, capturing potential false negatives of the screening process. No study of the M-CHAT or ITC adequately followed a large enough random sample of children to be able to realistically comment on negative predictive value (NPV), specificity, or sensitivity. Thoroughly following a population cohort of low risk children for a disorder with a low prevalence rate (1.47%) is a substantial challenge that requires significant resources and may be unrealistic for specific research programs. Specifically, case confirmation ultimately necessitates either expert behavioral assessment at young ages and/or thorough review of system records to determine case-ness or lack thereof. Current studies have not yet had the resources to provide such numbers. Without such studies, screening properties are evaluated by comparing PPV with the estimated prevalence of the disorder, and in some studies by assessing a select group of screen negatives, typically individuals who failed one portion of the screener or whose results were borderline.

Third, the available studies have high rates of attrition and there is very little information available about the children and families who leave the studies.

Fourth, data are lacking regarding how performance characteristics of screeners in primary care settings may vary based on child (e.g., sibling status, sex, prematurity, age, symptom severity and presentation) and family risk factors (e.g., SES, literacy level, race/ethnicity, primary language). Screening at ages below 16 months may result in more classification errors, specifically false positive screens for ASD in absence of other developmental concerns. Further, available data do not suggest an identifiable bias in terms of identification of children with more severe profiles. Specifically, developmental assessment of children within protocol confirmation samples identifies children with a range of cognitive, adaptive, and ASD symptom profiles. It is also important to note that such measurements are highly variable at young ages in that many children testing in ranges of impairment at very young ages may show substantial improvements via intervention and maturation at later ages.

Fifth, the majority of treatment studies are in clinically referred children, often older than the recommended screening age, and often with more severe manifestations of autism than would be expected in a screen detected population.

Finally, in terms of access to services, children screened at risk and later diagnosed with ASD receive higher levels of service in comparison to children without such diagnosis. However, systematic prospective studies have not yet been conducted that examine service access and the impact of service delivery post-diagnosis for children screened at-risk.

Future Research Needs

As noted, despite the resource-intensiveness and difficulty of execution, studies following large samples of both screen positive and negative participants would provide valuable data for understanding screening outcomes. Similarly, understanding of longer-term health outcomes of screening is largely lacking, and data on intermediate outcomes including time to diagnosis and treatment is sparse. Studies assessing potential modifiers of outcomes (age, sex, race, etc.) are also needed, as are studies evaluating specific risk factors such as prematurity and sibling status. Further, movement from screening, through diagnosis, to active engagement in evidence-based intervention is a complicated process inclusive of challenging child, family and system factors. Future work elucidating the specific characteristics of children, families, and systems that are more and less effectively engaged in the process is needed to address barriers on all levels. This would include evaluating the costs and benefits of the population level identification process that result from any attempts at screening. Largely, studies have investigated the ability to accurately detect the disorder early, but have less effectively understood those missed by this process and those where initial concerns were present, but seen as consistent with another area of concern or lack thereof. Such work would help both individuals and systems of care better weigh the benefits and harms of large scale screening programs. Importantly, the issues that affect families in whom children receive early false positive screens for ASD are not described in the literature and certainly warrant further consideration.

Comparative implementation of different screening practices and tools (ideally using randomized designs) including repeated screening procedures, combined use of measurement strategies (e.g., clinical observation/assessment), and head-to-head comparisons of tools would be valuable. Development of improved tools that may be better able to isolate critical concerning behaviors (at even earlier ages) across reporters as well as methods for understanding and accounting for reporting bias may also be productive areas of future work. Further, methodologies for combining biological and behavioral risk may be appropriate over time and help identify ASD at even earlier ages. Finally, studies to assess the effectiveness of treatment in screen detected populations are needed.

Conclusion

Formal ASD screening in general pediatric practices with the M-CHAT with follow-up interview, the M-CHAT-R/F procedure, has positive predictive values for identifying children with ASD of around 50 percent. Very little information is available on the degree to which screening tools miss positive cases, and no studies directly relate screening to clinical outcomes. Screening tools are widely available and referral pathways to early intervention and special education systems are available in all states of the U.S. The ability of individual families to access appropriate diagnostic and treatment resources within and across these settings is variable. Eleven of 20 early intensive intervention studies assessing cognitive changes demonstrated that treatment groups had statistically significant greater gains in cognitive outcomes than eclectic comparison groups. In 13 of 28 studies evaluating language outcomes and in nine of 10 evaluating joint attention outcomes, treatment was associated with significant improvements in outcomes relative to controls. Studies of long-term outcomes in populations with research identify early cognitive skills and shifts in cognitive outcomes as significant predictors of positive outcomes (including independence) at age 19.

References

1. Developmental Disabilities Monitoring Network Surveillance Year Principal I. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. MMWR Surveill Summ 2014 Mar 28;63 Suppl 2:1-21. PMID: 24670961.
2. Carbone PS, Murphy NA, Norlin C, et al. Parent and pediatrician perspectives regarding the primary care of children with autism spectrum disorders. J Autism Dev Disord 2013 Apr;43(4):964-72. PMID: 22948548.
3. Zuckerman KE, Mattox K, Donelan K, et al. Pediatrician identification of Latino children at risk for autism spectrum disorder. Pediatrics 2013 Sep;132(3):445-53. PMID: 23958770.
4. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. IV ed. Washington DC: American Psychiatric Association 2000.
5. Zwaigenbaum L, Bryson S, Lord C, et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. Pediatrics 2009 May;123(5):1383-91. PMID: 19403506.
6. Myers SM, Johnson CP. Management of children with autism spectrum disorders. Pediatrics 2007 Nov;120(5):1162-82. PMID: 17967921.
7. Myers SM. Management of autism spectrum disorders in primary care. Pediatr Ann 2009 Jan;38(1):42-9. PMID: 19213293.
8. American Psychiatric Association. DSM-5: The Future of Psychiatric Diagnosis. 2012. http://www.dsm5.org/Pages/Default.aspx.
9. Manning-Courtney P, Murray D, Currans K, et al. Autism spectrum disorders. Curr Probl Pediatr Adolesc Health Care 2013 Jan;43(1):2-11. PMID: 23332397.
10. Ozonoff S, Goodlin-Jones BL, Solomon M. Evidence-based assessment of autism spectrum disorders in children and adolescents. J Clin Child Adolesc Psychol 2005 Sep;34(3):523-40. PMID: 16083393.
11. Fecteau S, Mottron L, Berthiaume C, et al. Developmental changes of autistic symptoms. Autism 2003 Sep;7(3):255-68. PMID: 14516059.
12. Shattuck PT, Seltzer MM, Greenberg JS, et al. Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. J Autism Dev Disord 2007 Oct;37(9):1735-47. PMID: 17146700.
13. McGovern CW, Sigman M. Continuity and Change from Early Childhood to Adolescence in Autism. Journal of Child Psychology and Psychiatry 2005;46(4):401-8.
14. Seltzer MM, Shattuck P, Abbeduto L, et al. Trajectory of development in adolescents and adults with autism. Ment Retard Dev Disabil Res Rev 2004;10(4):234-47. PMID: 15666341.
15. Mahjouri S, Lord CE. What the DSM-5 Portends for Research, Diagnosis, and Treatment of Autism Spectrum Disorders. Curr Psychiatry Rep 2012 Dec;14(6):739-47. PMID: 22991100.
16. Billstedt E, Gillberg C, Gillberg C. Autism after Adolescence: Population-Based 13- to 22-Year Follow-Up Study of 120 Individuals with Autism Diagnosed in Childhood. Journal of Autism and Developmental Disorders 2005 Jun;35(3):351-60.

17. Eaves LC, Ho HH. Young Adult Outcome of Autism Spectrum Disorders. Journal of Autism and Developmental Disorders 2008 Apr;38(4):739-47.
18. Howlin P, Goode S, Hutton J, et al. Adult outcome for children with autism. J Child Psychol Psychiatry 2004 Feb;45(2):212-29. PMID: 14982237.
19. Taylor JL, Seltzer MM. Employment and post-secondary educational activities for young adults with autism spectrum disorders during the transition to adulthood. J Autism Dev Disord 2011 May;41(5):566-74. PMID: 20640591.
20. Folstein SE, Rosen-Sheidley B. Genetics of autism: complex aetiology for a heterogeneous disorder. Nat Rev Genet 2001 Dec;2(12):943-55. PMID: 11733747.
21. Bailey A, Le Couteur A, Gottesman I, et al. Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 1995 Jan;25(1):63-77. PMID: 7792363.
22. Marshall CR, Scherer SW. Detection and characterization of copy number variation in autism spectrum disorder. Methods Mol Biol 2012;838:115-35. PMID: 22228009.
23. Geschwind DH. Genetics of autism spectrum disorders. Trends Cogn Sci 2011 Sep;15(9):409-16. PMID: 21855394.
24. Hallmayer J, Cleveland S, Torres A, et al. Genetic heritability and shared environmental factors among twin pairs with autism. Arch Gen Psychiatry 2011 Nov;68(11):1095-102. PMID: 21727249.
25. Bristol MM, Cohen DJ, Costello EJ, et al. State of the science in autism: report to the National Institutes Health. J Autism Dev Disord 1996 Apr;26(2):121-54. PMID: 8744475.
26. Krakowiak P, Walker CK, Bremer AA, et al. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. Pediatrics 2012 May;129(5):e1121-8. PMID: 22492772.
27. Dodds L, Fell DB, Shea S, et al. The role of prenatal, obstetric and neonatal factors in the development of autism. J Autism Dev Disord 2011 Jul;41(7):891-902. PMID: 20922473.
28. Sandin S, Hultman CM, Kolevzon A, et al. Advancing maternal age is associated with increasing risk for autism: a review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2012 May;51(5):477-86 e1. PMID: 22525954.
29. Shelton JF, Tancredi DJ, Hertz-Picciotto I. Independent and dependent contributions of advanced maternal and paternal ages to autism risk. Autism Res 2010 Feb;3(1):30-9. PMID: 20143326.
30. Bhandari A, Sandlow JI, Brannigan RE. Risks to offspring associated with advanced paternal age. J Androl 2011 Mar-Apr;32(2):121-2. PMID: 20467047.
31. Gunnes N, Suren P, Bresnahan M, et al. Interpregnancy interval and risk of autistic disorder. Epidemiology 2013 Nov;24(6):906-12. PMID: 24045716.
32. Cheslack-Postava K, Liu K, Bearman PS. Closely spaced pregnancies are associated with increased odds of autism in California sibling births. Pediatrics 2011 Feb;127(2):246-53. PMID: 21220394.
33. Shelton JF, Hertz-Picciotto I, Pessah IN. Tipping the balance of autism risk: potential mechanisms linking pesticides and autism. Environ Health Perspect 2012 Jul;120(7):944-51. PMID: 22534084.
34. Schultz ST. Does thimerosal or other mercury exposure increase the risk for autism? A review of current literature. Acta Neurobiol Exp (Wars) 2010;70(2):187-95. PMID: 20628442.

35. Ozonoff S, Young GS, Carter A, et al. Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. Pediatrics 2011 Sep;128(3):e488-95. PMID: 21844053.
36. Gronborg TK, Schendel DE, Parner ET. Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. JAMA Pediatr 2013 Oct;167(10):947-53. PMID: 23959427.
37. Mandell DS, Listerud J, Levy SE, et al. Race differences in the age at diagnosis among medicaid-eligible children with autism. J Am Acad Child Adolesc Psychiatry 2002 Dec;41(12):1447-53. PMID: 12447031.
38. Warren ZE, Stone W. Best practices: Early diagnosis and psychological assessment. In: Autism Spectrum Disorders. New York: Oxford University Press; 2011:1271-82.
39. Pierce K, Carter C, Weinfeld M, et al. Detecting, studying, and treating autism early: the one-year well-baby check-up approach. J Pediatr 2011 Sep;159(3):458-65 e1-6. PMID: 21524759.
40. Stahmer AC, Mandell DS. State infant/toddler program policies for eligibility and services provision for young children with autism. Adm Policy Ment Health 2007 Jan;34(1):29-37. PMID: 16758329.
41. National Autism Center. National Standards Report. Randolph, MA: National Autism Center; 2009. Available from: http://www.nationalautismcenter.org/pdf/NAC%20Standards%20Report.pdf.
42. Committee on Educational Interventions for Children with Autism. Educating Children with Autism. Washington DC: National Academies Press; 2001. Available from: http://www.nap.edu/openbook.php?record_id=10017&page=R2.
43. Filipek PA, Accardo PJ, Ashwal S, et al. Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology 2000 Aug 22;55(4):468-79. PMID: 10953176.
44. Volkmar F, Cook E, Jr., Pomeroy J, et al. Summary of the Practice Parameters for the Assessment and Treatment of Children, Adolescents, and Adults with Autism and other Pervasive Developmental Disorders. American Academy of Child and Adolescent Psychiatry. J Am Acad Child Adolesc Psychiatry 1999 Dec;38(12):1611-6. PMID: 10596264.
45. Johnson CP, Myers SM. Identification and evaluation of children with autism spectrum disorders. Pediatrics 2007 Nov;120(5):1183-215. PMID: 17967920.
46. Technical report: the pediatrician's role in the diagnosis and management of autistic spectrum disorder in children. Pediatrics 2001 May;107(5):E85. PMID: 11331735.
47. UK National Screening Committee. The UK NSC recommendation on Autism screening in children. 2012. Available from: http://www.screening.nhs.uk/autism.
48. Scottish Intercollegiate Guidelines Network. Assessment, diagnosis and clinical interventions for children and young people with autism spectrum disorders. 2007. Available from: http://www.sign.ac.uk/pdf/sign98.pdf.
49. Swanson AR, Warren ZE, Stone WL, et al. The diagnosis of autism in community pediatric settings: Does advanced training facilitate practice change? Autism 2013 Jul 11PMID: 23847130.

50. Reed P, Osborne LA. Diagnostic practice and its impacts on parental health and child behaviour problems in autism spectrum disorders. Arch Dis Child 2012 Oct;97(10):927-31. PMID: 22806234.
51. Hyman SL, Johnson JK. Autism and Pediatric Practice: Toward a Medical Home. Journal of Autism and Developmental Disorders 2012;42(6):1156-64.
52. Dosreis S, Weiner CL, Johnson L, et al. Autism spectrum disorder screening and management practices among general pediatric providers. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S88-94. PMID: 16685190.
53. Kennedy T, Regehr G, Rosenfield J, et al. Exploring the gap between knowledge and behavior: a qualitative study of clinician action following an educational intervention. Acad Med 2004 May;79(5):386-93. PMID: 15107277.
54. Zwaigenbaum L, Bryson S, Garon N. Early identification of autism spectrum disorders. Behav Brain Res 2013 Aug 15;251:133-46. PMID: 23588272.
55. Warren Z, Veenstra-VanderWeele J, Stone W, Bruzek JL, Nahmias AS, Foss-Feig JH, Jerome RN, Krishnaswami S, Sathe NA, Glasser AM, Surawicz T, McPheeters ML. Therapies for Children With Autism Spectrum Disorders. Comparative Effectiveness Review No. 26. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I.) AHRQ Publication No. 11-EHC029-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2011. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm. PMID: 21834171.
56. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009 Oct;62(10):1013-20. PMID: 19230606.
57. U.S. Preventive Services Task Force. U.S. Preventive Services Task Force (USPSTF) Procedure Manual 2008. Available from: http://www.uspreventiveservicestaskforce.org/Page/Name/procedure-manual
58. Ben-Sasson A, Habib S, Tirosh E. Feasibility and Validity of Early Screening for Identifying Infants With Poor Social-Communication Development in a Well-Baby Clinic System. J Pediatr Nurs 2013 Nov 25PMID: 24333238.
59. Kamio Y, Inada N, Koyama T, et al. Effectiveness of Using the Modified Checklist for Autism in Toddlers in Two-Stage Screening of Autism Spectrum Disorder at the 18-Month Health Check-Up in Japan. J Autism Dev Disord 2013 Jun 6PMID: 23740200.
60. Posserud M-B, Breivik K, Gillberg C, et al. ASSERT—The Autism Symptom SElf-ReporT for adolescents and adults: Bifactor analysis and validation in a large adolescent population. Research in Developmental Disabilities 2013;34(12):4495-503.
61. Robins DL, Casagrande K, Barton M, et al. Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). Pediatrics 2013 Dec 23PMID: 24366990.
62. Nygren G, Sandberg E, Gillstedt F, et al. A new screening programme for autism in a general population of Swedish toddlers. Res Dev Disabil 2012 Jul-Aug;33(4):1200-10. PMID: 22502846.
63. Miller JS, Gabrielsen T, Villalobos M, et al. The each child study: systematic screening for autism spectrum disorders in a pediatric setting. Pediatrics 2011 May;127(5):866-71. PMID: 21482605.

64. VanDenHeuvel A, Fitzgerald M, Greiner B, et al. Screening for autistic spectrum disorder at the 18-month developmental assessment: a population-based study. Ir Med J 2007 Sep;100(8):565-7. PMID: 17955715.
65. Dietz C, Swinkels SH, van Daalen E, et al. Parental compliance after screening social development in toddlers. Arch Pediatr Adolesc Med 2007 Apr;161(4):363-8. PMID: 17404133.
66. Dietz C, Swinkels S, van Daalen E, et al. Screening for autistic spectrum disorder in children aged 14-15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. J Autism Dev Disord 2006 Aug;36(6):713-22. PMID: 16633887.
67. Kleinman JM, Robins DL, Ventola PE, et al. The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. J Autism Dev Disord 2008 May;38(5):827-39. PMID: 17882539.
68. Bi C, Wu J, Jiang T, et al. Mutations of ANK3 identified by exome sequencing are associated with autism susceptibility. Hum Mutat 2012 Dec;33(12):1635-8. PMID: 22865819.
69. Robins DL. Screening for autism spectrum disorders in primary care settings. Autism 2008 Sep;12(5):537-56. PMID: 18805946.
70. Pandey J, Verbalis A, Robins DL, et al. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. Autism 2008 Sep;12(5):513-35. PMID: 18805945.
71. Robins DL, Fein D, Barton ML, et al. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. J Autism Dev Disord 2001 Apr;31(2):131-44. PMID: 11450812.
72. Honda H, Shimizu Y, Nitto Y, et al. Extraction and Refinement Strategy for detection of autism in 18-month-olds: a guarantee of higher sensitivity and specificity in the process of mass screening. J Child Psychol Psychiatry 2009 Aug;50(8):972-81. PMID: 19298465.
73. Barbaro J, Ridgway L, Dissanayake C. Developmental surveillance of infants and toddlers by maternal and child health nurses in an Australian community-based setting: promoting the early identification of autism spectrum disorders. J Pediatr Nurs 2011 Aug;26(4):334-47. PMID: 21726784.
74. Barbaro J, Dissanayake C. Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: the social attention and communication study. J Dev Behav Pediatr 2010 Jun;31(5):376-85. PMID: 20495475.
75. Baird G, Charman T, Baron-Cohen S, et al. A screening instrument for autism at 18 months of age: a 6-year follow-up study. J Am Acad Child Adolesc Psychiatry 2000 Jun;39(6):694-702. PMID: 10846303.
76. Wetherby AM, Brosnan-Maddox S, Peace V, et al. Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. Autism 2008;12(5):487-511.
77. Inada N, Koyama T, Inokuchi E, et al. Reliability and validity of the Japanese version of the Modified Checklist for autism in toddlers (M-CHAT). Research in Autism Spectrum Disorders 2011;5(1):330-6.
78. Canal-Bedia R, Garcia-Primo P, Martin-Cilleros MV, et al. Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. J Autism Dev Disord 2011 Oct;41(10):1342-51. PMID: 21161677.

79. Turner-Brown LM, Baranek GT, Reznick JS, et al. The First Year Inventory: A longitudinal follow-up of 12-month-old to 3-year-old children. Autism 2013;17(5):527-40.
80. Oosterling I, Rommelse N, de Jonge M, et al. How useful is the Social Communication Questionnaire in toddlers at risk of autism spectrum disorder? J Child Psychol Psychiatry 2010 Nov;51(11):1260-8. PMID: 20626528.
81. Oosterling IJ, Swinkels SH, van der Gaag RJ, et al. Comparative analysis of three screening instruments for autism spectrum disorder in toddlers at high risk. J Autism Dev Disord 2009 Jun;39(6):897-909. PMID: 19205862.
82. Oosterling IJ, Wensing M, Swinkels SH, et al. Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. J Child Psychol Psychiatry 2010 Mar;51(3):250-8. PMID: 19843319.
83. Koyama T, Inokuchi E, Inada N, et al. Utility of the Japanese version of the Checklist for Autism in Toddlers for predicting pervasive developmental disorders at age 2. Psychiatry Clin Neurosci 2010 Jun;64(3):330-2. PMID: 20602732.
84. LoVullo SV, Matson JL. Development of a critical item algorithm for the Baby and Infant Screen for Children with aUtIsm Traits. Research in Autism Spectrum Disorders 2012;6(1):378-84.
85. Kozlowski AM, Matson JL, Worley JA, et al. Defining characteristics for young children meeting cutoff on the Modified Checklist for Autism in Toddlers. Research in Autism Spectrum Disorders 2012;6(1):472-9.
86. Matson JL, Wilkins J, Sharp B, et al. Sensitivity and specificity of the Baby and Infant Screen for Children with Autism Traits (BISCUIT): Validity and cutoff scores for autism and PDD-NOS in toddlers. Research in Autism Spectrum Disorders 2009;3(4):924-30.
87. Matson JL, Fodstad JC, Dempsey T. What symptoms predict the diagnosis of autism or PDD-NOS in infants and toddlers with developmental delays using the Baby and Infant Screen for aUtIsm Traits. Dev Neurorehabil 2009;12(6):381-8. PMID: 20205546.
88. Gray KM, Tonge BJ, Sweeney DJ, et al. Screening for autism in young children with developmental delay: an evaluation of the developmental behaviour checklist: early screen. J Autism Dev Disord 2008 Jul;38(6):1003-10. PMID: 17972169.
89. Corsello C, Hus V, Pickles A, et al. Between a ROC and a hard place: decision making and making decisions about using the SCQ. J Child Psychol Psychiatry 2007 Sep;48(9):932-40. PMID: 17714378.
90. Nadig AS, Ozonoff S, Young GS, et al. A prospective study of response to name in infants at risk for autism. Arch Pediatr Adolesc Med 2007 Apr;161(4):378-83. PMID: 17404135.
91. Allen CW, Silove N, Williams K, et al. Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. J Autism Dev Disord 2007 Aug;37(7):1272-8. PMID: 17080270.
92. Eaves LC, Wingert H, Ho HH. Screening for autism: agreement with diagnosis. Autism 2006 May;10(3):229-42. PMID: 16682396.
93. Eaves LC, Ho HH. The very early identification of autism: outcome to age 4 1/2-5. J Autism Dev Disord 2004 Aug;34(4):367-78. PMID: 15449513.
94. Feldman MA, Ward RA, Savona D, et al. Development and initial validation of a parent report measure of the behavioral development of infants at risk for autism spectrum disorders. J Autism Dev Disord 2012 Jan;42(1):13-22. PMID: 21387111.

95. Snow AV, Lecavalier L. Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. Autism 2008 Nov;12(6):627-44. PMID: 19005032.
96. Myers CL, Gross AD, McReynolds BM. Broadband Behavior Rating Scales as Screeners for Autism? J Autism Dev Disord 2013 Nov 28. PMID: 24287879.
97. Oner P, Oner O, Munir K. Three-item Direct Observation Screen (TIDOS) for autism spectrum disorder. Autism 2013 Oct 14. PMID: 24126869.
98. Koh HC, Lim SH, Chan GJ, et al. The Clinical Utility of the Modified Checklist for Autism in Toddlers with High Risk 18-48 Month Old Children in Singapore. J Autism Dev Disord 2013 Jun 28. PMID: 23812662.
99. Williams K, Perkins D, Wheeler D, et al. Can questions about social interaction correctly identify preschool aged children with autism? J Paediatr Child Health 2013 Feb;49(2):E167-74. PMID: 23350819.
100. Narzisi A, Calderoni S, Maestro S, et al. Child Behavior Check List 1(1/2)-5 as a tool to identify toddlers with autism spectrum disorders: a case-control study. Res Dev Disabil 2013 Apr;34(4):1179-89. PMID: 23376628.
101. Chlebowski C, Robins DL, Barton ML, et al. Large-scale use of the modified checklist for autism in low-risk toddlers. Pediatrics 2013 Apr;131(4):e1121-7. PMID: 23530174.
102. Kamio Y, Inada N, Koyama T. A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders. Autism 2013;17(1):15-26.
103. Dawson G, Rogers S, Munson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics 2010 Jan;125(1):e17-23. PMID: 19948568.
104. Dawson G, Jones EJ, Merkle K, et al. Early behavioral intervention is associated with normalized brain activity in young children with autism. J Am Acad Child Adolesc Psychiatry 2012 Nov;51(11):1150-9. PMID: 23101741.
105. Itzchak EB, Zachor DA. Who Benefits from Early Intervention in Autism Spectrum Disorders? Research in Autism Spectrum Disorders 2011;5(1):345-50.
106. Zachor DA, Itzchak EB. Treatment Approach, Autism Severity and Intervention Outcomes in Young Children. Research in Autism Spectrum Disorders 2010;4(3):425-32.
107. Eldevik S, Hastings RP, Jahr E, et al. Outcomes of behavioral intervention for children with autism in mainstream pre-school settings. J Autism Dev Disord 2012 Feb;42(2):210-20. PMID: 21472360.
108. Eikeseth S, Klintwall L, Jahr E, et al. Outcome for Children with Autism Receiving Early and Intensive Behavioral Intervention in Mainstream Preschool and Kindergarten Settings. Research in Autism Spectrum Disorders 2012;6(2):829-35.
109. Flanagan HE, Perry A, Freeman NL. Effectiveness of large-scale community-based intensive Behavioral Intervention: A waitlist comparison study exploring outcomes and predictors. Research in Autism Spectrum Disorders 2012;6(2):673-82.
110. Zachor DA, Ben-Itzchak E, Rabinovich A-L, et al. Change in autism core symptoms with intervention. Research in Autism Spectrum Disorders 2007;1(4):304-17.
111. Howard JS, Sparkman CR, Cohen HG, et al. A comparison of intensive behavior analytic and eclectic treatments for young children with autism. Res Dev Disabil 2005 Jul-Aug;26(4):359-83. PMID: 15766629.

112. Strain PS, Bovey EH. Randomized, Controlled Trial of the Leap Model of Early Intervention for Young Children with Autism Spectrum Disorders. Topics in Early Childhood Special Education 2011;31(3):133-54.
113. Boyd BA, Hume K, McBee MT, et al. Comparative Efficacy of LEAP, TEACCH and Non-Model-Specific Special Education Programs for Preschoolers with Autism Spectrum Disorders. J Autism Dev Disord 2013 Jun 28 PMID: 23812661.
114. Peters-Scheffer N, Didden R, Mulders M, et al. Low intensity behavioral treatment supplementing preschool services for young children with autism spectrum disorders and severe to mild intellectual disability. Res Dev Disabil 2010 Nov-Dec;31(6):1678-84. PMID: 20627451.
115. Hayward D, Eikeseth S, Gale C, et al. Assessing progress during treatment for young children with autism receiving intensive behavioural interventions. Autism 2009 Nov;13(6):613-33. PMID: 19933766.
116. Eikeseth S, Hayward D, Gale C, et al. Intensity of Supervision and Outcome for Preschool Aged Children Receiving Early and Intensive Behavioral Interventions: A Preliminary Study. Research in Autism Spectrum Disorders 2009 Jan;3(1):67-73.
117. Eikeseth S, Smith T, Jahr E, et al. Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. Behav Modif 2002 Jan;26(1):49-68. PMID: 11799654.
118. Eikeseth S, Smith T, Jahr E, et al. Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: a comparison controlled study. Behav Modif 2007 May;31(3):264-78. PMID: 17438342.
119. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S145-55. PMID: 16685181.
120. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. Am J Ment Retard 2000 Jul;105(4):269-85. PMID: 10934569.
121. Peters-Scheffer N, Didden R, Mulders M, et al. Effectiveness of low intensity behavioral treatment for children with autism spectrum disorder and intellectual disability. Research in Autism Spectrum Disorders 2013;7(9):1012-25.
122. Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four-year outcome and predictors. Am J Ment Retard 2005 Nov;110(6):417-38. PMID: 16212446.
123. Reed P, Osborne LA, Corness M. Brief report: relative effectiveness of different home-based behavioral approaches to early teaching intervention. J Autism Dev Disord 2007 Oct;37(9):1815-21. PMID: 17180714.
124. Landa RJ, Holman KC, O'Neill AH, et al. Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: a randomized controlled trial. J Child Psychol Psychiatry 2011 Jan;52(1):13-21. PMID: 21126245.
125. Landa RJ, Kalb LG. Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention. Pediatrics 2012 Nov;130 Suppl 2:S186-90. PMID: 23118250.

126. Strauss K, Vicari S, Valeri G, et al. Parent inclusion in Early Intensive Behavioral Intervention: the influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. Res Dev Disabil 2012 Mar-Apr;33(2):688-703. PMID: 22188793.
127. Fava L, Strauss K, Valeri G, et al. The Effectiveness of a Cross-Setting Complementary Staff- and Parent-Mediated Early Intensive Behavioral Intervention for Young Children with ASD. Research in Autism Spectrum Disorders 2011;5(4):1479-92. PMID: 881460693; EJ927555.
128. Casenhiser DM, Shanker SG, Stieben J. Learning through interaction in children with autism: preliminary data from asocial-communication-based intervention. Autism 2013 Mar;17(2):220-41. PMID: 21949005.
129. Rogers SJ, Estes A, Lord C, et al. Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. Journal of the American Academy of Child & Adolescent Psychiatry 2012;51(10):1052-65.
130. Estes A, Vismara L, Mercado C, et al. The Impact of Parent-Delivered Intervention on Parents of Very Young Children with Autism. J Autism Dev Disord 2013 Jul 10 PMID: 23838727.
131. Roberts J, Williams K, Carter M, et al. A Randomised Controlled Trial of Two Early Intervention Programs for Young Children with Autism: Centre-Based with Parent Program and Home-Based. Research in Autism Spectrum Disorders 2011;5(4):1553-66.
132. Reed P, Osborne L. Impact of Severity of Autism and Intervention Time-Input on Child Outcomes: Comparison across Several Early Interventions. British Journal of Special Education 2012;39(3):130-6.
133. Oosterling I, Visser J, Swinkels S, et al. Randomized controlled trial of the focus parent training for toddlers with autism: 1-year outcome. J Autism Dev Disord 2010 Dec;40(12):1447-58. PMID: 20440639.
134. Carter AS, Messinger DS, Stone WL, et al. A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. Journal of Child Psychology and Psychiatry 2011;52(7):741-52.
135. Green J, Charman T, McConachie H, et al. Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. Lancet 2010 May 20 PMID: 20494434.
136. Drew A, Baird G, Baron-Cohen S, et al. A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. Eur Child Adolesc Psychiatry 2002 Dec;11(6):266-72. PMID: 12541005.
137. Aldred C, Green J, Adams C. A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness. J Child Psychol Psychiatry 2004 Nov;45(8):1420-30. PMID: 15482502.
138. Aldred C, Green J, Emsley R, et al. Mediation of treatment effect in a communication intervention for pre-school children with autism. Journal of Autism and Developmental Disorders 2012;42(3):447-54.
139. Keen D, Couzens D, Muspratt S, et al. The effects of a parent-focused intervention for children with a recent diagnosis of autism spectrum disorder on parenting stress and competence. Research in Autism Spectrum Disorders 2010;4(2):229-41.

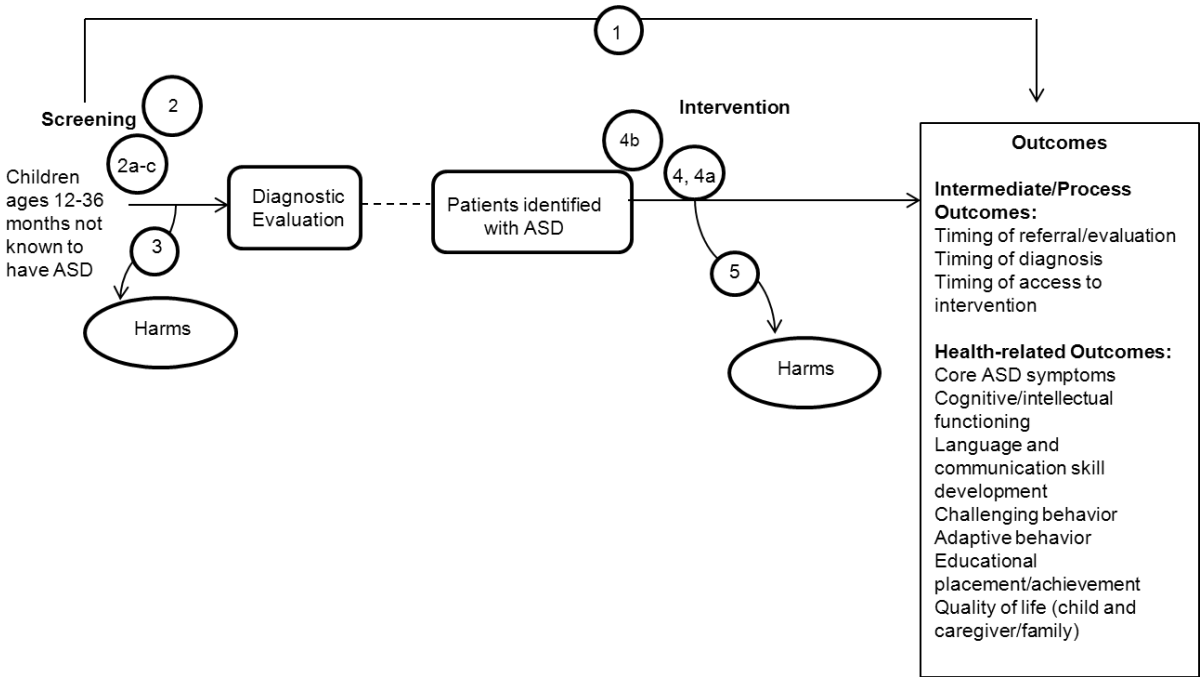
140. Schreibman L, Stahmer AC. A Randomized Trial Comparison of the Effects of Verbal and Pictorial Naturalistic Communication Strategies on Spoken Language for Young Children with Autism. J Autism Dev Disord 2013 Nov 23 PMID: 24272416.
141. Perry A, Blacklock K, Dunn Geier J. The relative importance of age and IQ as predictors of outcomes in Intensive Behavioral Intervention. Research in Autism Spectrum Disorders 2013;7(9):1142-50.
142. Kasari C, Lawton K, Shih W, et al. Caregiver-mediated intervention for low-resourced preschoolers with autism: an RCT. Pediatrics 2014 Jul;134(1):e72-9. PMID: 24958585.
143. Wong CS. A play and joint attention intervention for teachers of young children with Autism: A randomized controlled pilot study. Autism 2013 May 2013;17(3):340-57.
144. Warreyn P, Roeyers H. See what I see, do as I do: Promoting joint attention and imitation in preschoolers with autism spectrum disorder. Autism 2013 Oct 8 PMID: 24104513.
145. Venker CE, McDuffie A, Weismer SE, et al. Increasing Verbal Responsiveness in Parents of Children with Autism: A Pilot Study. Autism: The International Journal of Research and Practice 2012;16(6):568-85.
146. Kaale A, Smith L, Sponheim E. A randomized controlled trial of preschool-based joint attention intervention for children with autism. J Child Psychol Psychiatry 2012 Jan;53(1):97-105. PMID: 21883204.
147. Siller M, Hutman T, Sigman M. A Parent-Mediated Intervention to Increase Responsive Parental Behaviors and Child Communication in Children with ASD: A Randomized Clinical Trial. Journal of Autism and Developmental Disorders 2013;43(3):540-55.
148. Lawton K, Kasari C. Teacher-implemented joint attention intervention: pilot randomized controlled study for preschoolers with autism. J Consult Clin Psychol 2012 Aug;80(4):687-93. PMID: 22582764.
149. Kasari C, Freeman S, Paparella T. Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. J Child Psychol Psychiatry 2006 Jun;47(6):611-20. PMID: 16712638.
150. Kasari C, Gulsrud A, Freeman S, et al. Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. J Am Acad Child Adolesc Psychiatry 2012 May;51(5):487-95. PMID: 22525955.
151. Kasari C, Paparella T, Freeman S, et al. Language outcome in autism: randomized comparison of joint attention and play interventions. J Consult Clin Psychol 2008 Feb;76(1):125-37. PMID: 18229990.
152. Lawton K, Kasari C. Brief Report: Longitudinal Improvements in the Quality of Joint Attention in Preschool Children with Autism. Journal of Autism and Developmental Disorders 2012;42(2):307-12.
153. Schertz HH, Odom SL, Baggett KM, et al. Effects of Joint Attention Mediated Learning for Toddlers with Autism Spectrum Disorders: An Initial Randomized Controlled Study. Early Childhood Research Quarterly 2013 2013;28(2):249-58.
154. Kasari C, Gulsrud AC, Wong C, et al. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. Journal of Autism and Developmental Disorders 2010;40(9):1045-56.
155. Goods KS, Ishijima E, Chang Y-C, et al. Preschool Based JASPER Intervention in Minimally Verbal Children with Autism: Pilot RCT. Journal of Autism and Developmental Disorders 2013 May 2013;43(5):1050-6.

156. Ingersoll B. Brief report: Pilot randomized controlled trial of reciprocal imitation training for teaching elicited and spontaneous imitation to children with autism. Journal of Autism and Developmental Disorders 2010;40(9):1154-60.
157. Ingersoll B. Brief report: effect of a focused imitation intervention on social functioning in children with autism. J Autism Dev Disord 2012 Aug;42(8):1768-73. PMID: 22146934.
158. Pinto-Martin JA, Levy SE, Feldman JF, et al. Prevalence of autism spectrum disorder in adolescents born weighing <2000 grams. Pediatrics 2011 Nov;128(5):883-91. PMID: 22007018.
159. Stephens BE, Bann CM, Watson VE, et al. Screening for autism spectrum disorders in extremely preterm infants. J Dev Behav Pediatr 2012 Sep;33(7):535-41. PMID: 22926660.
;160. Limperopoulos C, Bassan H, Sullivan NR, et al. Positive screening for autism in ex-preterm infants: prevalence and risk factors. Pediatrics 2008 Apr;121(4):758-65. PMID: 18381541.
161. Kuban KC, O'Shea TM, Allred EN, et al. Positive screening on the Modified Checklist for Autism in Toddlers (M-CHAT) in extremely low gestational age newborns. J Pediatr 2009 Apr;154(4):535-40 e1. PMID: 19185317.
162. Guevara JP, Gerdes M, Localio R, et al. Effectiveness of developmental screening in an urban setting. Pediatrics 2013 Jan;131(1):30-7. PMID: 23248223.
163. Anderson DK, Liang JW, Lord C. Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. J Child Psychol Psychiatry 2013 Dec 9PMID: 24313878.
164. Dawson G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol 2008 Summer;20(3):775-803. PMID: 18606031.
165. Poon KK, Watson LR, Baranek GT, et al. To what extent do joint attention, imitation, and object play behaviors in infancy predict later communication and intellectual functioning in ASD? J Autism Dev Disord 2012 Jun;42(6):1064-74. PMID: 21858585.
166. Koegel LK, Koegel RL, Carter CM. Pivotal responses and the natural language teaching paradigm. Semin Speech Lang 1998;19(4):355-71; quiz 72; 424. PMID: 9857392.
167. Reichow B, Barton EE, Boyd BA, et al. Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). Cochrane Database Syst Rev 2012;10:CD009260. PMID: 23076956.
168. Peters-Scheffer N, Didden R, Korzilius H, et al. A Meta-Analytic Study on the Effectiveness of Comprehensive ABA-Based Early Intervention Programs for Children with Autism Spectrum Disorders. Research in Autism Spectrum Disorders 2011;5(1):60-9.
169. Virues-Ortega J. Applied behavior analytic intervention for autism in early childhood: meta-analysis, meta-regression and dose-response meta-analysis of multiple outcomes. Clin Psychol Rev 2010 Jun;30(4):387-99. PMID: 20223569.
170. Reichow B. Overview of Meta-Analyses on Early Intensive Behavioral Intervention for Young Children with Autism Spectrum Disorders. Journal of Autism and Developmental Disorders 2012;42(4):512-20.

171. Kuppens S, Onghena P. Sequential Meta-Analysis to Determine the Sufficiency of Cumulative Knowledge: The Case of Early Intensive Behavioral Intervention for Children with Autism Spectrum Disorders. Research in Autism Spectrum Disorders 2012;6(1):168-76.
172. South Carolina Department of Disabilities and Special Needs. South Carolina Department of Disabilities and Special Needs. Available at: http://ddsn.sc.gov/Pages/default.aspx
173. Zwaigenbaum L. Assessment and diagnosis of autism spectrum disorders. BMJ 2011;343:d6628. PMID: 22021469.
174. Warren Z, Stone W, Humberd Q. A training model for the diagnosis of autism in community pediatric practice. J Dev Behav Pediatr 2009 Oct;30(5):442-6. PMID: 19823138.
175. Autism Speaks A. Autism Speaks: State initiatives. Available at: http://www.autismspeaks.org/advocacy/states.
176. Luyster R, Gotham K, Guthrie W, et al. The Autism Diagnostic Observation Schedule-toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. J Autism Dev Disord 2009 Sep;39(9):1305-20. PMID: 19415479.
177. Guthrie W, Swineford LB, Nottke C, et al. Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation. J Child Psychol Psychiatry 2013 May;54(5):582-90. PMID: 23078094.
178. Charman T, Taylor E, Drew A, et al. Outcome at 7 years of children diagnosed with autism at age 2: predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. J Child Psychol Psychiatry 2005 May;46(5):500-13. PMID: 15845130.
179. Risi S, Lord C, Gotham K, et al. Combining Information from Multiple Sources in the Diagnosis of Autism Spectrum Disorders. Journal of the American Academy of Child and Adolescent Psychiatry 2006;45(9):1094-103.
180. Stone WL, Lee EB, Ashford L, et al. Can autism be diagnosed accurately in children under 3 years? J Child Psychol Psychiatry 1999 Feb;40(2):219-26. PMID: 10188704.
181. Landa R, Garrett-Mayer E. Development in Infants with Autism Spectrum Disorders: A Prospective Study. Journal of Child Psychology and Psychiatry 2006;47(6):629-38.
182. Chawarska K, Klin A, Paul R, et al. Autism spectrum disorder in the second year: stability and change in syndrome expression. J Child Psychol Psychiatry 2007 Feb;48(2):128-38. PMID: 17300551.
183. Corsello CM, Akshoomoff N, Stahmer AC. Diagnosis of autism spectrum disorders in 2-year-olds: a study of community practice. J Child Psychol Psychiatry 2013 Feb;54(2):178-85. PMID: 22905987.
184. Fein D, Barton M, Eigsti IM, et al. Optimal outcome in individuals with a history of autism. J Child Psychol Psychiatry 2013 Feb;54(2):195-205. PMID: 23320807.
185. Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. J Consult Clin Psychol 1987 Feb;55(1):3-9. PMID: 3571656.
186. Rutter M. Autistic children: infancy to adulthood. Semin Psychiatry 1970 Nov;2(4):435-50. PMID: 5527410.
187. Orinstein AJ, Helt M, Troyb E, et al. Intervention for optimal outcome in children and adolescents with a history of autism. J Dev Behav Pediatr 2014 May;35(4):247-56. PMID: 24799263.

188. Magiati I, Tay XW, Howlin P. Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: a systematic review of longitudinal follow-up studies in adulthood. Clin Psychol Rev 2014 Feb;34(1):73-86. PMID: 24424351.
189. Howlin P, Moss P, Savage S, et al. Outcomes in Adult Life Among Siblings of Individuals with Autism. J Autism Dev Disord 2014 Sep 5 PMID: 25189825.
190. Shattuck PT, Durkin M, Maenner M, et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. J Am Acad Child Adolesc Psychiatry 2009 May;48(5):474-83. PMID: 19318992.
191. Goin-Kochel RP, Mackintosh VH, Myers BJ. How many doctors does it take to make an autism spectrum diagnosis? Autism 2006 Sep;10(5):439-51. PMID: 16940311.
192. Fountain C, King MD, Bearman PS. Age of diagnosis for autism: individual and community factors across 10 birth cohorts. J Epidemiol Community Health 2011 Jun;65(6):503-10. PMID: 20974836.
193. Mandell DS, Ittenbach RF, Levy SE, et al. Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. J Autism Dev Disord 2007 Oct;37(9):1795-802. PMID: 17160456.
194. Mandell DS, Wiggins LD, Carpenter LA, et al. Racial/ethnic disparities in the identification of children with autism spectrum disorders. Am J Public Health 2009 Mar;99(3):493-8. PMID: 19106426.

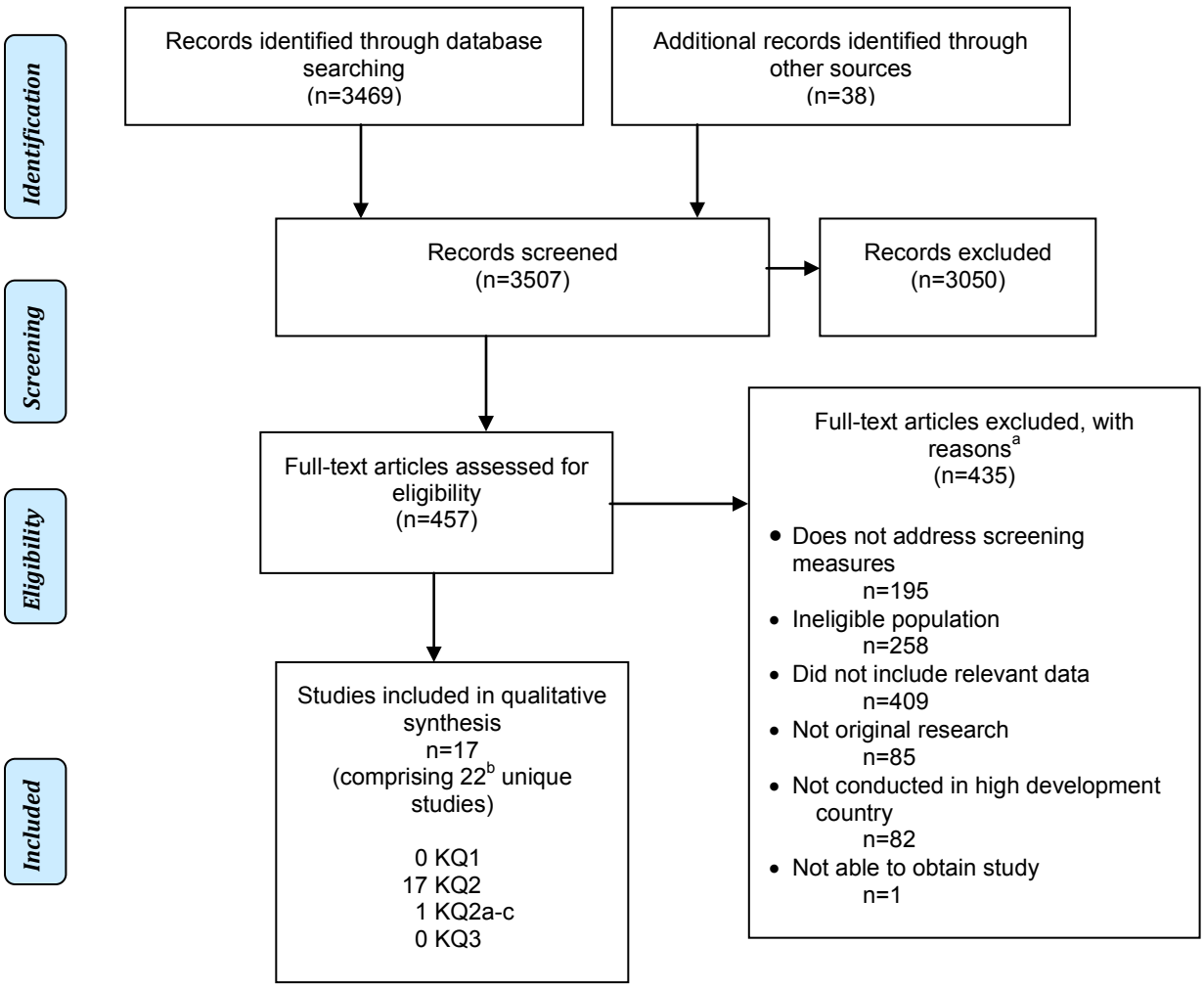
Figure 1. Analytic Framework



Note: Numbers in circles on the diagram refer to key questions.

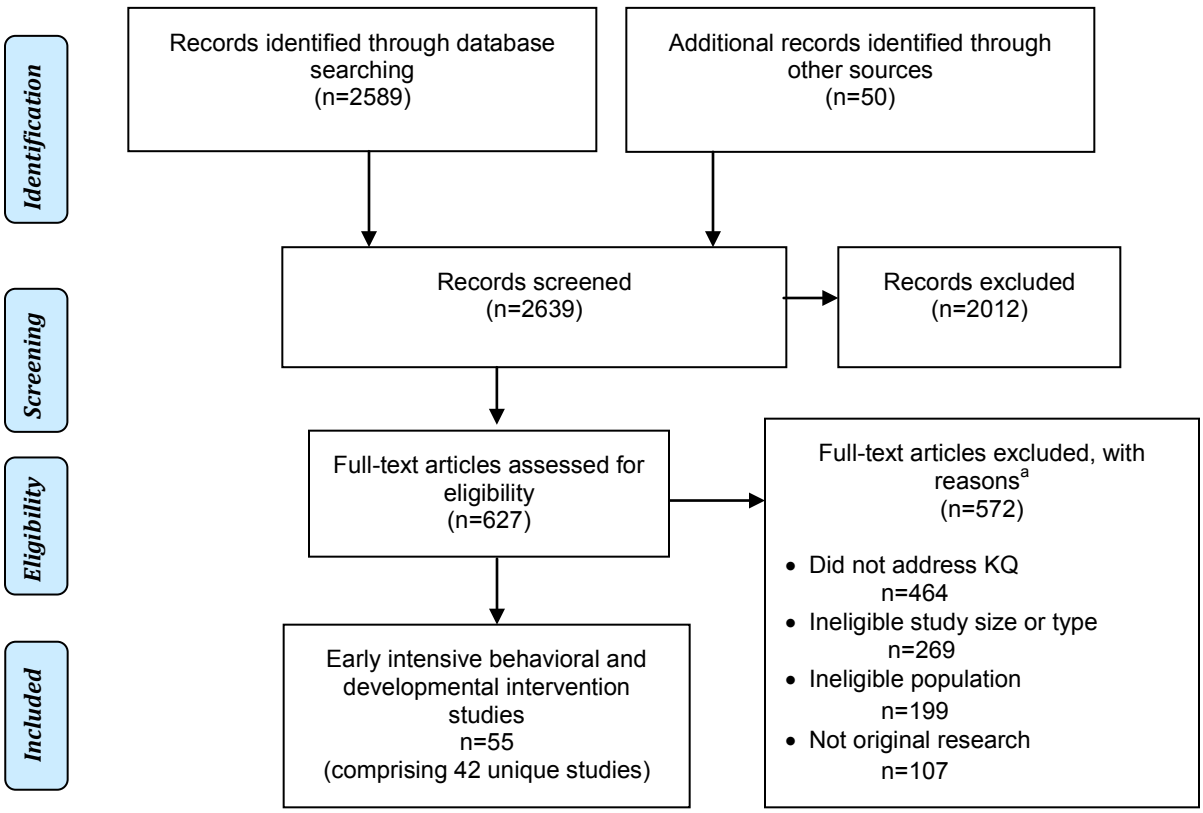
Abbreviation: ASD=autism spectrum disorder.

Figure 2. Disposition of Screening Studies Identified for This Review (KQs 1–3)



[a] Numbers do not tally as studies could be excluded for multiple reasons.
[b] One paper reported 2 unique studies; 3 studies comprised multiple publications.

Figure 3. Disposition of Intervention Studies Identified for This Review (KQ 4)



[a] Numbers do not tally as studies could be excluded for multiple reasons.

Box 1. Commonly Used ASD Screening Tools

Screening Tool	Format	Administration Time	Key Characteristics
Checklist for Autism in Toddlers (CHAT; 1992)	Parent-rated scale + Professional observation	5 minutes	• Targets children at 18 months old • Assesses communication, joint attention, pretend play using yes/no questions • First screener systematically assessing early ASD identification • No longer in common use
Modified Checklist for Autism in Toddlers (M-CHAT) with and without M-CHAT Follow-up interview (M-CHAT/F; 2001)	Parent-rated scale + Follow-up interview	5-10 minutes	• Targets children 16-30 months old • Assesses communication, joint attention, pretend play • Includes follow-up interview that was used in most but not all studies.
Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F; 2014) *Revision of M-CHAT/F*	Parent-rated scale + Follow-up interview procedure	5 minutes	• Targets young children 16-30 months of age • Assesses communication, joint attention, pretend play • Offers explicit follow-up procedures based on scores from both instrument and follow-up interview
Infant Toddler Checklist (ITC; 2002)	Parent-rated scale+ parent follow-up interview	5-10 minutes	• Targets children 9-24 months old • Screener has available norms across developmental ages for risk
Early Screening of Autistic Traits (ESAT; 2006)	Parent-rated scale + clinician observation	5-10 minutes	• Measure originally developed in Dutch and used primarily in European studies • Targets children 14-15 months old • Requires pre-screening in clinical setting followed by in-home parent-completed questionnaire
First Year Inventory/First Year Inventory-Lite (FYI-L; 2003)	Parent rated scale	5-10 minutes	• Targets children at 12 months of age • Assesses skills regarding communication, joint attention, socialization, and sensory regulation
Social Attention and Communication Study Checklist (SACS; 2010)	Clinician rated scale	Unknown	• Designed for repeated assessment of children at ages 8, 12, 18, and 24 months old • Trained nursing providers conduct behavioral observations during informal interactions • Measures markers of social interaction, play skills, and verbal and nonverbal communication
Young Autism and other developmental disorders Checkup Tool (YACHT; 2009)	Parent-rated scale, + Professional observation + Child Performance	5-20 Minutes	• Japanese measure targeting 18 month old children • Assesses early social communication markers via parent-report, clinician observation, and brief testing

Abbreviation: ASD=autism spectrum disorder.

Box 2. Characteristics of Included Screening Studies

Study, Country, Quality	Instrument (Cut Point for Positive Screen)	Age Range	Total Screened	Group(s) Followed to Diagnosis
Baird 2000[75] UK Good	CHAT; first screen (NR)	18.7±1.1 months	16,235	Attempted to follow all, but had substantial loss to followup
VanDenHeuvel 2007[64] Ireland Poor	CHAT + followup CHAT for participants at risk on first screen (NR)	18-20 months	2117	Screen positive
Robins 2014[61] US Good	M-CHAT-R/F (≥2)	16-31 months	16,115	Screen positive, 42 negative screen STAT positive and/or pediatric concerns
Chlebowski 2013[101] US Good	M-CHAT + followup interview (3/23 items failed or 2/6 critical items)	Mean, 20 months	18,989	Screen positive
Kamio 2014[59] Japan Fair	M-CHAT-Japanese + followup interview	17-26 months	1851	All
Nygren 2012[62] Sweden Fair	M-CHAT + followup interview (3/23 items failed or 2/6 critical items)	2.5 years	3999	Screen positive
Canal-Bedia 2011[78] Stage 1 Spain Fair	M-CHAT (Spanish) + followup interview (3/23 items failed or 2/6 critical items)	18-24 months	2480	Screen positive
Canal-Bedia 2011[78] Stage 2 Spain Fair	M-CHAT (Spanish) + followup interview (3/23 items failed or 2/6 critical items)	18-24 months	2055	Screen positive
Inada 2011[77] Japan Fair-Poor	M-CHAT-J; Full (2/23 items failed)	4-26 months	1187	Unclear; published analysis includes only children with both 18 and 36 month data
Pierce 2011[39] US Good	CSBS-ITC (10th percentile)	10.1-15.9 months	10,479	Screen positive participants seen for ≥2 sessions and 41 randomly selected screen negative participants
Wetherby 2008[76] US Good	ITC (bottom 10th percentile on social or symbolic composite or total score or bottom 10th percentile on 2 consecutive speech composite)	6-24 months	5385	Screen positive + random screen negative participants (number NR)
Miller 2011[63] US Fair	M-CHAT + followup interview (3/23 items failed or 2/6 critical items) and ITC (bottom 10th percentile)	14-32 months	796	Screen positive and 2 screen negative participants
Turner-Brown 2014[79] US Fair	FYI (≥90th percentile)	12 months	1305	Screen positive
Ben-Sasson 2013[58] Israel Fair	FYI	11-13 months	583	Screen positive + 12 screen negative + review of medical records of 148 screen negatives
Dietz 2007[65,66] Netherlands Fair	ESAT-4 item (≥3 items failed)	14 – 15 months	31724	Screen positive
Barbaro 2011[73,74] Australia Fair	SACS (3/4 or 5 items failed)	8-24 months	20,770	Screen positive
Honda 2009[72] Japan Fair	YACHT-18 (NR)	18 months	2814	Screen positive + 4 screen negative participants

Box 2. Characteristics of Included Screening Studies

Abbreviations: ASD=autism spectrum disorder; CHAT=Checklist for Autism in Toddlers; ESAT=Early Screen for Autistic Traits; ITC=Infant Toddler Checklist; JA-OBS=Joint Attention Observation Schedule; M-CHAT=Modified Checklist for Autism in Toddlers; M-CHAT-R/F=Modified Checklist for Autism in Toddlers, Revised with Follow-up; PPV=positive predictive value; SACS=Social Attention and Communication Study; YACHT=Young Autism and other Developmental Disorders Check-Up Tool.

Table 1. Inclusion and Exclusion Criteria

	Include	Exclude
Definition of Disease	Clinical diagnosis of autism spectrum disorder (ASD)	
Populations		
KQs 1-3	Young children ages 12 to 36 months undergoing screening for ASD	Studies that exclusively focus on infants or older children or adults or that assess general developmental screening
KQs 4, 5	Young children ages 0-12 years undergoing intervention for ASD	Studies of treatments for adolescents or adults
Interventions		
KQ 1	Tools and approaches used specifically to screen for ASD	Studies of screening for other conditions, general developmental screening or genetic or biomarker screening
KQ 2	Tools and approaches used specifically to screen for ASD	Studies of screening for other conditions, general developmental screening or genetic or biomarker screening
KQ 3	Tools and approaches used specifically to screen for ASD	Studies of screening for other conditions, general developmental screening or genetic or biomarker screening
KQs 4, 5	Interventions for ASD directed to children 0-5 years	Studies of interventions directed to older children, adolescents or adults as these would not follow from screening; medication interventions not relevant to young children
Comparators		
KQs 1-3	No screening or alternate screening approaches when comparing two or more approaches	
KQs 4,5	Placebo, other intervention, no intervention	Noncomparative studies
Outcomes		
KQ 1	Timing of referral/evaluation, diagnosis, access to intervention; core ASD symptoms; cognitive and intellectual functioning; language/communication skill development; challenging behavior; adaptive behavior; educational placement/achievement; quality of life for child and family assessed in preschool and elementary school age groups	Short term outcomes such as changes in joint attention
KQ 2	Performance characteristics of ASD screening approaches	Studies that do not allow calculation of performance characteristics
KQ 3	Harms from ASD screening (e.g., labeling, family distress)	
KQ 4	Core ASD symptoms, cognitive and intellectual functioning, language and communication skill development, challenging behavior, adaptive behavior, educational placement/achievement, and quality of life for the child and family	Short term outcomes such as changes in joint attention
KQ 5	Harms of ASD treatment (e.g., worsening of behavior, other unintended consequences)	
Setting		
KQs 1-3	Primary care settings and primary care referable settings, early intervention and education settings	Studies not conducted in one of the following countries, rated as “very high” on the United Nations’ International Human Development Index (HDI): Andorra, Argentina, Australia, Austria, Bahrain, Barbados, Belgium, Brunei Darussalam, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong (China), Hungary, Iceland, Ireland, Israel, Italy, Japan, Republic of Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, New

Table 1. Inclusion and Exclusion Criteria

	Include	Exclude
		Zealand, Norway, Poland, Portugal, Qatar, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States
KQs 4, 5	Clinical, educational, early intervention, or home settings	
Study Designs		
KQs 1-3 (benefits or harms of screening, screening performance)	Any study design except single case reports	Single case reports
KQs 4, 5 (benefits or harms of intervention)	Any comparative study design targeting young children with ASD, including at least 10 children with ASD; Good quality systematic reviews	Single case reports Studies with <10 children with ASD Studies included in prior comprehensive systematic review

Abbreviations: ASD=autism spectrum disorder; KQ=key question.

Table 2. Quality Scoring Algorithm for Intervention Studies

Definition and Scoring Algorithm	Rating
Score algorithm for internal validity quality rating	
≥8/10 points, including a ++ on study design and ++ on diagnostic approach	Good quality
≥6/10 points, including at least a + on intervention	Fair quality
≤5/10 points	Poor quality

Table 3. Performance Characteristics of Screeners in Primary Care/Unselected Samples Without Accounting for Attrition

Study, Country, Quality Age Range	Instrument (Cut Point for Positive Screen)	Step 1: Total Screen Positive/Total Screened LFU From Step 1 to Step 2, n (%)	Step 2: Total Screen Positive/ Total Screened LFU From Screening to Diagnosis, n (%)	Diagnostic Workup: Total Diagnosed With ASD via the Screening Tool/Total Evaluated	% Identified With ASD via Screening Rate Identified With ASD per 1000	Study PPV (%)	Other Available Data
Baird 2000[75] UK Good Age: 18.7 ± 1.1 months	CHAT (NR)	407/16235 347 (85.3)	34/60 0	20/34	0.12% 2 per 1000	58.8	Se: 35.1 Sp: 97.7 PPV for any developmental issue: 18.9%
VanDenHeuvel 2007[64] Ireland Poor Age: 18-20 months	CHAT+followup CHAT for participants at risk on first screen (NR)	29/2117 10 (34.5)	7/19 0	7/12	0.33% 3 per 1000	58.3	PPV for any developmental issue: 75%
Robins 2014[61] US Good Age: 16-30 months	M-CHAT-R/F (≥2)	1155/16115 209/1155	348/946 156/946 (16.5)	105/221	0.65% 6.5 per 1000	48	PPV for any developmental issue: 95%
Chlebowski 2013[101] US Good Age: 20 months	M-CHAT+ followup interview (3/23 items failed or 2/6 critical items)	1737/18989 442 /1737 (25.4)	272/1295 107/272 (39.3)	98/207	0.48% 5 per 1000	54	PPV for any developmental issue: 97.5%
Kamio 2014[59] Japan Fair Age: 17-26 months	M-CHAT+ followup-Japanese (3/23 items failed or 1 or 2/6 critical items)	319/1851 124/319 (38.9)	44/195 0	42/1727	1.08% 11 per 1000	45.5	PPV for any developmental issue: 54.5%
Nygren 2012[62] Sweden Fair Age: 2.5 years	M-CHAT+followup interview (3/23 items failed or 2/6 critical items)	36/3999 NR		33/49	0.83% 8 per 1000	91.7	Se: 76.7%
Canal-Bedia 2011[78] Stage 1 Spain Fair Age: 18-24 months	M-CHAT (Spanish) +followup interview (3/23 items failed or 2/6 critical items)	429/2480 NA	- 343/429 (80)	23/86	0.93% 9 per 1000	26.7	PPV for any developmental issue: 50%
Canal-Bedia 2011[78] Stage 2 Spain Fair Age: 18-24 months	M-CHAT (Spanish) +followup interview (3/23 items failed or 2/6 critical items)	336/2055 NA	- 305/336 (90.8)	6/31	0.29% 2.9 per 1000	19	NR

Table 3. Performance Characteristics of Screeners in Primary Care/Unselected Samples Without Accounting for Attrition

Study, Country, Quality Age Range	Instrument (Cut Point for Positive Screen)	Step 1: Total Screen Positive/Total Screened LFU From Step 1 to Step 2, n (%)	Step 2: Total Screen Positive/ Total Screened LFU From Screening to Diagnosis, n (%)	Diagnostic Workup: Total Diagnosed With ASD via the Screening Tool/Total Evaluated	% Identified With ASD via Screening Rate Identified With ASD per 1000	Study PPV (%)	Other Available Data
Inada 2011[77] Japan Fair-Poor Age: 4-26 months	M-CHAT-J Full (2 of any 23 items failed)	NR/1187	NR	NR	1.68% 17 per 1000	10.7 (data to re-calculate not provided)	Predicted: PPV: 10.7 Se: 75 Sp: 89 NPV: 99.5
Pierce 2011[39] US Good Age: 10.1 -15.9 months	CSBS-ITC (10th percentile)	1318/10,479 NA	- 1134/1318 (86)	32/184	0.31% 3 per 1000	17.4	PPV for any developmental issue: 72.3%
Wetherby 2008[76] US Good Age: 6-24 months	ITC (bottom 10th percentile on social or symbolic composite or total score or bottom 10th percentile on 2 consecutive speech composite)	482/5385 NA	- 422/482 (87.6)	56/60	1.04% 10 per 1000	NR	NR
Miller 2011[63] US Fair Age: 14-32 months	ITC (bottom 10th percentile) or M-CHAT+followup Interview (3/23 items failed or 2/6 critical items)	47/796 0		13/27	1.38% 14 per 1000	32	PPV for any developmental issue: 96.4%
Ben-Sasson 2013[58] Israel Fair Age: 11-13 months	FYI (total score 0.42)	15/583	NA	NR	NR	60 (social commun-ication delay)	NR
Turner-Brown 2014[79] US Fair Age: 12 months	FYI (≥90th percentile)	64/1305 NR	- 46/64 (71.9)	4/699	0.31% 3 per 1000	NR	NR
Dietz 2007[65,66] Netherlands Fair Age: 14-15 months	ESAT-4 item (≥3 items failed)	370/31,724 115/370 (31)	100/255 34/100 (34)	19/66	0.06% 0.6 per 1000	29	NR

Table 3. Performance Characteristics of Screeners in Primary Care/Unselected Samples Without Accounting for Attrition

Study, Country, Quality Age Range	Instrument (Cut Point for Positive Screen)	Step 1: Total Screen Positive/Total Screened LFU From Step 1 to Step 2, n (%)	Step 2: Total Screen Positive/ Total Screened LFU From Screening to Diagnosis, n (%)	Diagnostic Workup: Total Diagnosed With ASD via the Screening Tool/Total Evaluated	% Identified With ASD via Screening Rate Identified With ASD per 1000	Study PPV (%)	Other Available Data
Barbaro 2011[73,74] Australia Fair Age: 8-24 months	SACS (3/4 or 5 items failed)	216/20,770 NR	- 106/216 (49.1)	89/110	0.43% 4.3 per 1000	81 (24 months)	PPV for any developmental issue: 99%
Honda 2009[72] Japan Fair Age: 18 months	YACHT-18 (NR)	402/2814 NR	- 379/402 (94.3)	14/23	0.39% 3.9 per 1000	58	NR

Abbreviations: ASD=autism spectrum disorder; CHAT=Checklist for Autism in Toddlers; ESAT=Early Screen for Autistic Traits; ITC=Infant Toddler Checklist; JA-OBS=Joint Attention Observation Schedule; LFU=loss to followup; M-CHAT=Modified Checklist for Autism in Toddlers; PPV=positive predictive value; SACS=Social Attention and Communication Study; Se=sensitivity; Sp=specificity; YACHT=Young Autism and Other Developmental Disorders Check-Up Tool.

Table 4. PPV of Screeners in Primary Care/Unselected Samples Without Accounting for Attrition

Study, Country, Quality Age Range	Instrument (Cut Point for Positive Screen) Time to Administer	Number Screening Positive at Step 1/Total Number Screened at Step 1 (%)	Number Screening Positive at Step 2/Total Number Screened at Step 2 (%)	PPV for ASD, %	N False-Positive	N (%) Diagnosed With ASD	Rate of ASD
Robins 2014[61] US Good 16-30 months	M-CHAT-R/F (≥2) 5 minutes	1155/16,115 (7.2)	348/946 (37)	48	116	123 (46.8)	6.5 per 1000
Chlebowski 2013[101] US Good Mean, 20 months	M-CHAT+followup interview (3/23 items failed or 2/6 critical items) 5-10 minutes	1737/18,989 (9.1)	272/1295 (21)	54	79	98 (47.3)	5 per 1000
Kamio 2014[59] Japan Fair 17-26 months	M-CHAT-Japanese (3/23 items failed or 1 or 2/6 critical items) 5-10 minutes	319/1851 (17.2)	44/195 (22.6)	45.5	24	42 (2.43)	11 per 1000
Nygren 2012[62] Sweden Fair 2.5 years	M-CHAT+followup interview (3/23 items failed or 2/6 critical items) 5-10 minutes	36/3999 (0.9)	NA	91.7	3	33 (67.3)	8 per 1000
Canal-Bedia 2011[78] Stage 1 Spain Fair 18-24 months	M-CHAT (Spanish)+ followup interview (3/23 items failed or 2/6 critical items) 5-10 minutes	429/2480 (17.2)	NA	26.7	63	23 (26.7)	9 per 1000
Canal-Bedia 2011[78] Stage 2 Spain Fair 18-24 months	M-CHAT (Spanish)+ followup interview (3/23 items failed or 2/6 critical items) 5-10 minutes	336/2055 (16.4)	NA	19	25	6 (19.4)	2.9 per 1000
Inada 2011[77] Japan Fair-Poor 4-26 months	M-CHAT-J Full (2 of any 23 items failed) 5-10 minutes	NR/1187	NR	10.7*	125*	20 (1.69)*	17 per 1000
Pierce 2011[39] US Good 10.1-15.9 months	CSBS-ITC (10th percentile) 5-10 minutes	1318/10,479 (12.6)	NA	17.4	152	32 (17.4)	3 per 1000

Table 4. PPV of Screeners in Primary Care/Unselected Samples Without Accounting for Attrition

Study, Country, Quality Age Range	Instrument (Cut Point for Positive Screen) Time to Administer	Number Screening Positive at Step 1/Total Number Screened at Step 1 (%)	Number Screening Positive at Step 2/Total Number Screened at Step 2 (%)	PPV for ASD, %	N False-Positive	N (%) Diagnosed With ASD	Rate of ASD
Wetherby 2008[76] US Good 6-24 months	ITC (bottom 10th percentile on social or symbolic composite or total score or bottom 10th percentile on 2 consecutive speech composite) 5-10 minutes	482/5385 (9)	NA	NR	NR	56 (93.3)	10 per 1000
Miller 2011[63] US Fair 14-32 months	ITC (bottom 10th percentile) or M-CHAT +followup interview (3/23 items failed or 2/6 critical items) 5-10 minutes	47/796 (5.9)	NA	32	17	13 (43.3)	14 per 1000

* Predicted.

Abbreviations: ASD=autism spectrum disorder; CHAT=Checklist for Autism in Toddlers; ITC=Infant Toddler Checklist; M-CHAT=Modified Checklist for Autism in Toddlers; PPV=positive predictive value.

Table 5. Performance Characteristics of Screeners in Large Primary Care/Unselected Samples (>15,000 Participants) Correcting for Attrition and Other Developmental Concerns Warranting Further Evaluation for Early Intervention

Study, Country, Quality	Instrument (Cut Point for Positive Screen)	Total Screen Positive/Total Screened	Diagnosed With ASD/Total Evaluated	% Identified With ASD Correcting for Attrition	Imputed Rate of ASD Correcting for Attrition
Chlebowski 2013[101] US Good	M-CHAT+followup interview (3/23 items failed or 2/6 critical items)	272/18,989	92/171	1.01	1 in 98
Robins, 2014[61] US Good	M-CHAT-R/F (≥2)	348/16,115	105/221	.92	1 in 109

Abbreviations: ASD=autism spectrum disorder; M-CHAT=Modified Checklist for Autism in Toddlers.

Table 6. Summary of Cognitive and Language Outcome Measures Frequently Used in Studies Included in the Review

Measure	Description/Purpose	Age Range for Administration	Total Score Range	Range for Abnormal Scores
British Ability Scales	Multi-scale tool to assess cognitive ability and educational achievement; includes vocabulary, pattern, reading, and recall tests	2.5 or 3 years to 17 years 11 months		High scores indicate more abilities.
Bayley Scales of Infant Development	Describes the current developmental functioning of infants and to assist in diagnosis and treatment planning for infants with developmental delays or disabilities. The test is intended to measure a child's level of development in three domains: cognitive, motor, and behavioral.	From 1 to 42 months of age	Mean=10, SD=3 (range from 1-19) Mean=100, SD=15 (range from 40-160)	Average: 90-109 Low Average: 89-89 Borderline: 70-79 Extremely Low: <69
Differential Abilities Scale	In-depth analysis of children's learning abilities. General Conceptual Ability (GCA): summarizes developmentally appropriate ability domains Special Nonverbal Composite (SNC): summarizes the nonverbal domains (the SNC is particularly valuable when testing children who are not proficient in spoken English)	Lower Preschool (ages 2 years, 6 months through 3 years, 5 months), Upper Preschool (aged 3 years, 6 months through 5 years, 11 months), and School-Age (6 years, 0 months through 17 years, 11 months)	Standard scores with mean of 100 (50th percentile) and SD of 15	Higher scores indicate more abilities. Average range: 85-115.
Developmental Assessment of Young Children	The five subtests relate to the areas of development that are evaluated for early intervention eligibility and children's developmental performance (Cognition, Communication, Social-Emotional, Physical Development, Adaptive Behavior)	Birth through 5 years and 11 months of age	0-1 (0=skill not observed; 1=skill was observed). Add up the 1s to calculate the raw score, then convert raw score to standard score based on age group	Higher scores indicate more skills shown.
Early Social Communication Scales	Structured assessment designed to provide measures of individual differences in nonverbal communication skills in children with mental ages between 8 and 30 months of age	8 months to 30 months	Codes frequency of social-communication behaviors based on developmental stage, communicative goal, and whether they were child-initiated.	Higher scores indicate more skills shown. Frequency counts are compared to means from a small standardization sample based upon the category, with values varying by subscale.
Gilliam Autism Rating Scale	Identifying autism in individuals and estimating its severity; six subscales: Restrictive/Repetitive Behaviors, Social Interaction, Social Communication, Emotional Responses, Cognitive Style, and Maladaptive Speech	3 to 22 years	Cutoff scores based on standardization sample	>85: Likely ASD 70-85: Possible ASD <70: Unlikely ASD Higher scores indicate more impairment.

Table 6. Summary of Cognitive and Language Outcome Measures Frequently Used in Studies Included in the Review

Measure	Description/Purpose	Age Range for Administration	Total Score Range	Range for Abnormal Scores
Griffiths Mental Development Scales	Measure the rate of development of young children related to motor, social, and reasoning skills	2 to 8 years	Ratio transformation: divide mental age by chronological age, which yields different means and SDs for each subscale. Means tend to be around 100 and SDs tend to be around 16, but this varies depending on the scale and the participant's age	Higher scores indicate more abilities.
MacArthur-Bates Communicative Development Inventories	Evaluate the communicative skills of young typically developing children from their "early signs of comprehension, to their first nonverbal gestural signals, to the expansion of early vocabulary and the beginnings of grammar"; Pt 1. Early words and Pt 2. Actions and gestures	8 to 18 months	Total words and gestures used are compared to maximum values within each domain and percentile ranks are calculated.	Percentile ranks are used to calculate deviation from the norm, with >2 SDs from the mean typically considered significantly impaired. Higher scores indicate more abilities.
Merrill-Palmer Scale of Mental Tests	Intelligence test that primarily assesses nonverbal skills. The scale evaluates both the content of thinking (that is, the material that is actually processed by the child) and the process of thinking (that is, the way in which this material is used to form new concepts).	18 months to four years	Mental age	Average: 85-115 Below Average: 70-84 Low: <70 Higher scores indicate more abilities.
Mullen Scales of Early Learning	Assesses early intellectual development and school readiness, permitting targeted intervention at a young age - Five scales: Gross Motor, Visual Reception, Fine Motor, Expressive Language, and Receptive Language.	Birth to 5 years, 8 months	Standard scores with mean of 100 (50^{th} percentile) and SD of 15.	Higher scores indicate more abilities. Average range: 85-115.
Psychoeducational Profile	Assesses skills and behaviors of children with autism and communication disabilities, identifying learning strengths, uneven development, emerging abilities, and other information useful in educational programming - Communication, Motor, and Maladaptive Behaviors	6 months through 7 years	Yields a total developmental score but inconsistently correlates with other IQ measures. All items within each subscale are summed to provide subscale and overall total scores.	Higher scores indicate more abilities.
Preschool Language Scale	Interactive assessment of developmental language skills. Total language, auditory comprehension, expressive communication standard scores, growth scores, percentile ranks, language age equivalents	Birth to 7 years, 11 months	Standard scores with mean of 100 (50^{th} percentile) and SD of 15.	Higher scores indicate more abilities. Average range: 85-115.
Peabody Picture Vocabulary Test	A measure of receptive vocabulary for Standard American English; verbal ability and scholastic aptitude	4-5 and 10-11 years of age	Standard scores with mean of 100 (50^{th} percentile) and SD of 15.	Higher scores indicate more abilities. Average range: 85-115.

Table 6. Summary of Cognitive and Language Outcome Measures Frequently Used in Studies Included in the Review

Measure	Description/Purpose	Age Range for Administration	Total Score Range	Range for Abnormal Scores
Reynell Developmental Language Scale	Standardized measure of language development. The Verbal Comprehension Scale measures language understanding of nouns, verbs, and prepositions and following of simple directions and more complex directions. The Verbal Comprehension Scale also provides information on early developing inferencing skills. The Expressive language Scale measures spoken language in the areas of syntax, vocabulary, and content.	1-6 years		Higher scores indicate more abilities.
Wechsler Intelligence Scale for Children-Revised, WISC-III, or WISC-IV	A general test of intelligence; defined as "the global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment." Full Scale IQ, Verbal IQ and Performance IQ	6 to 16 years 11 months	Mean = 100, SD = 15 Standard score of 100 = 50th percentile rank	Higher scores indicate more abilities. ≥130: Very Superior 120-129: Superior 110-119: High Average 90-109: Average 80-89: Low Average 70-79: Borderline ≤69: Extremely Low
Wechsler Preschool and Primary Scale of Intelligence, Revised	Intelligence test; The Verbal Scale measures language expression, comprehension, listening, and the ability to apply these skills to solving problems. The examiner gives the questions orally, and the child gives a spoken response. The Performance Scale assesses nonverbal problem solving, perceptual organization, speed, and visual-motor proficiency. Included are tasks like puzzles, analysis of pictures, imitating designs with blocks, and copying.	2 years, 6 months to 7 years, 3 months	Same as WISC	Same as WISC

Table 7. Overview of Intervention Studies Reporting Cognitive Outcomes

Intervention Approach	N Studies Reporting Cognitive Outcomes	Intervention>Control, n Studies	No Significant Between-Group Differences, n Studies
Direct to Child			
UCLA/Lovaas	8	4	4
ESDM	1	1	0
LEAP	1	1	0
Preschool-delivered models	5	4	1
Parent Training			
Parent-delivered ESDM	1	0	1
AEPS	1	0	1
Clinic and/or home-based	2	1	1
Parent training in communication responsiveness	1	0	1
Total	**20**	**11**	**9**

Intervention>Control=outcomes were significantly improved in the intervention vs. comparison arm.

Table 8. Overview of Intervention Studies Reporting Language Outcomes

Intervention Approach	n Reporting Language Outcomes	Intervention>Control, n Studies	No Significant Between-Group Differences, n Studies
Direct to Child			
UCLA/Lovaas	7	3	4
ESDM	1	1	0
LEAP	2	1	1
Preschool-delivered models	6	5	1
Parent Training			
Parent-delivered ESDM	1	0	1
AEPS	1	0	1
Clinic and/or home-based	2	2	0
Parent training in communication responsiveness	5	1	4
DIR/Floortime	1	0	1
More than Words	1	0	1
Pivotal response training	1	0	1
Total	**28**	**13**	**15**

Intervention>Control=outcomes were significantly improved in the intervention vs. comparison arm.

Table 9. Overview of Intervention Studies Reporting Joint Attention Outcomes

Intervention Approach	n Reporting Joint Attention Outcomes	Intervention>Control, n Studies	No Significant Between-Group Differences, n Studies
Play/interaction-based	10	9	1

Table 10. Summary of Cognitive Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
RCTs					
Strain et al, 2011[112] US **IG:** LEAP program with coaching and training, 28 classrooms (27 analyzed)/177 children **CG:** LEAP intervention manuals only, 28 classrooms (23 analyzed)/117 children **Quality:** Fair	**IG:** 50.1±4.6 **CG:** 50.7±4.2	MSEL, composite	**IG:** 59.6±6.9 **CG:** 63.2±6.6	**IG:** 68.6±7.5 **CG:** 61.4±9.0	IG change: 0.89 CG change: -1.8 Intervention group had MSEL-Early Learning Composite scores significantly higher than controls at 2 years, with an effect size of 0.89, range: 0.59 to 1.22, p<0.01
Dawson et al, 2012[103,104] US **IG:** ESDM, 24/24 **CG:** Eclectic community-based interventions, 24/21 **Quality:** Good	**IG:** 23.9±4.0 **CG:** 23.1±3.9	MSEL, composite	**IG:** 61±9.2 **CG:** 59.4±8.6	**IG:** 78.6±24.2 **CG:** 66.3±15.3	At 2 years, MSEL composite increased by 17.6 in the treated group versus 7.0 in the comparator Group x Time (baseline vs. 2 year): F=4.31, p=0.044
Sallows et al, 2005[122] US **IG:** Clinic directed UCLA/Lovaas-based early intensive intervention, 13/13 **CG:** Parent-directed UCLA/Lovaas-based early intensive intervention, 10/10 **Quality:** Good	**IG:** 33.23±3.89 **CG:** 34.20±5.06	BSID, MPSMT, Full IQ	**IG:** 50.85±10.57 **CG:** 52.10±8.98	**IG:** 73.08±33.08 **CG:** 79.60±21.80	No significant differences between groups at pre- or posttest
		Non-Verbal IQ	**IG:** 70.58±16.54 **CG:** 82.67±14.94	**IG:** 77.58±25.24 **CG:** 89.44±18.35	
Smith et al, 2000[120] US **IG:** UCLA/Lovaas-based intervention, 15/15 **CG:** Parent training from Lovaas manual, 13/13 **Quality:** Fair	**IG:** 36.07±6.00 **CG:** 35.77±5.77	Stanford Binet-IQ	**IG:** 50.53±11.18 **CG:** 50.69±13.88	**IG:** 66.49±24.08 **CG:** 49.67±19.74	Increase in IQ in IG significantly greater than comparison group; using pooled variance t tests, for group differences, p<0.05
		PSMT	**IG:** 21.60±4.49 **CG:** 21.92±5.5	**IG:** 64.33±18.74 **CG:** 49.17±21.43	

Table 10. Summary of Cognitive Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Non-randomized trials					
Peters-Scheffer et al, 2013[121] Netherlands **IG:** Low intensity Lovaas-based intervention + specialized preschool, 20/20 **CG:** Specialized preschool, 20/20 **Quality:** Good	**IG+CG:** 62.52 ± 16.96	MSEL	**IG:** 23.34 (7.32) **CG:** 23.43 (6.34)	At 24 months: **IG:** 39.70 (11.99) **CG:** 32.44 (11.55)	Significantly greater effects in IG compared to preschool in developmental age (d=1.09, p=0.001) and IQ (d=0.40, p<0.001)
		Developmental age, in months	**IG:** 40.66 (20.07) **CG:** 40.14 (18.27)	**IG:** 48.12 (19.71) **CG:** 39.42 (19.89)	
Peters-Scheffer et al, 2010[114] Netherlands **IG:** Specialized preschool +UCLA/Lovaas-based intervention, 12/12 **CG:** Eclectic preschool, 22/22 **Quality:** Fair	**IG:** 53.5±5.52 **CG:** 52.95±11.14	WPPSI-R, BSID, Snijders-Oomen Nonverbal Intelligence Test-Revised	Developmental age in months **IG:** 25.92±7.57 **CG:** 23.32±6.33	Developmental age in months **IG:** 34.83±10.89 **CG:** 25.73±8.26	GLM repeated measures assessed group differences over time: Developmental Age: F(1,32)=23.37 Mental Development Index/IQ: F(1,32)=26.96
			Mental developmental index/IQ **IG:** 47.00±10.33 **CG:** 45.73±15.99	Mental developmental index/IQ **IG:** 55.83±14.94 **CG:** 43.73±16.74	After baseline equivalence established, significantly greater improvement in IG vs. CG, p<0.01
Hayward et al, 2009[115,116] UK **IG:** Intensive clinic-based UCLA/Lovaas-based intervention, 23/20 **CG:** Intensive parent-managed treatment, 21/19 **Quality:** Fair	**IG:** 35.7±6.2 **CG:** 34.4±5.7	BSID	**IG:** 53.5±15.1 **CG:** 54.1±15.1	**IG:** 70.9±19.6 **CG:** 68.9±22.1	Non-significant group differences both at intake and at followup (independent t-tests).
		WPPSI–R	**IG:** 74.8±22.6 **CG:** 76.2±18.2	**IG:** 89.4±29.2 **CG:** 82.1±28.0	
Reed et al, 2007[123] UK **IG:** High intensity intervention, 14/14 **IGa:** High intensity w/focus on Lovaas techniques, 4/4 **IGb:** High intensity w/focus on verbal behavior, 5/5 **IGc:** High intensity w/focus on CABAS methods, 5/5 **CG:** Low intensity intervention in home-based direct teaching sessions, 13/13 **Quality:** Fair	**IG:** 42.9 (14.8) **IGa:** 47.5 (13.5) **IGb:** 38.0 (9.9) **IGc:** 44.2 (20.5) **CG:** 40.8 (5.6)	PEP-R	**IG:** 57.2±17.8 **IGa:** 58.0±30.7 **IGb:** 50.2±7.7 **IGc:** 63.6±12.4 **CG:** 49.3±13.2	Overall score, mean change (ES): **IG:** NR **IGa:** NR (0.91) **IGb:** NR (0.82) **IGc:** NR (1.11) **CG:** NR	Change scores for educational functioning were statistically higher in the high intensity arms compared to the low intensity arm t(25)=2.54, p<0.01 between groups
		BAS, cognitive ability score	**IG:** 60.1±22.4 **IGa:** 72.0±30.6 **IGb:** 48.0±4.6 **IGc:** 62.8±23.9 **CG:** 52.4±9.9	BAS cognitive ability score, mean change (ES): **IG:** NR **IGa:** NR (0.58) **IGb:** NR (3.74) **IGc:** NR (3.74) **CG:** NR	

Table 10. Summary of Cognitive Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Eikeseth et al, 2002[117,118] Norway **IG:** UCLA/Lovaas-based intervention, 13/13 **CG:** Eclectic therapy, 12/12 **Quality:** Fair	**IG:** 66.31±14.71 **CG:** 65±10.95	WPPSI-R, WISC-R, BSID, MPSMT	IQ: **IG:** 61.92±11.31 **CG:** 65.17±14.97 Performance IQ: **IG:** 77.54±30.21 **CG:** 81.83±21.05	IQ: **IG:** 79.08±18.09 **CG:** 69.50±18.38 Performance IQ: **IG:** 95±16.91 **CG:** 90.17±19.97	Change scores: IQ: **IG:** 17.15±10.97 **CG:** 4.33±7.55 p<0.01 1.5 years later, differences were maintained, with significantly greater improvement in IQ in the Lovaas group (25 points) compared to eclectic therapy (7 points), p<0.05 Performance IQ: **IG:** 17.46±30.70 **CG:** 8.33±16.12 p=NS
Cohort studies					
Eldevik et al, 2012[107] Norway **IG:** Preschool-based early intensive intervention, 31/31 **CG:** Eclectic preschool, 12/12 **Quality:** Fair	**IG:** 42.2±9.0 **CG:** 46.2±12.4	BSID, Stanford-Binet, WPPSI-R	Intellectual functioning, mean±SD (range): **IG:** 51.6±16.9 (24-94) **CG:** 51.7±18.1 (30-89)	Intellectual functioning, mean±SD (range): **IG:** 66.6±24.8 (23-110) **CG:** 52.2±22.0 (23-86)	Mean change in intellectual functioning in treatment group was 15.1 (SD 14.9) compared to 0.5 (SD 9.5) in treatment as usual group
Flanagan et al, 2012[109,141] Canada **IG:** Intensive behavioral intervention, 61/61 **CG:** Eclectic interventions, 61/61 **Quality:** Fair	**IG:** 42.93±11.53 **CG:** 42.79±10.51	MSEL, WPPSI, Stanford-Binet	**IG:** NR **CG:** NR	**IG:** 55.80±26.97 **CG:** 39.50±18.93 Marginal mean scores: **IG:** 55.71 **CG:** 36.46	Higher cognitive scores in IG vs. CG at followup (19 point difference), p=0.003 Effect size=0.83
Itzchak et al, 2011[105,106] Israel **IG:** ABA-based approach, 45/45 **CG:** Eclectic approach, 33/33 **Quality:** Fair	**IG:** 25.1±3.9 **CG:** 26.0±4.6	MSEL Verbal	**IG+CG:** 60.9±24.4	**IG+CG:** 75.0±27	MANOVA repeated measures over time: F=422.3, n=63, p<0.001, ES=0.406
		Non-Verbal	**IG+CG:** 73.9±23.7	**IG+CG:** 75.5±29.2	No group difference in non-verbal scores: F=0.1, ES=0.001 (n=63)
Zachor et al, 2007[110] Israel **IG:** UCLA/Lovaas-based intervention, 53/53 **CG:** Eclectic approach, 15/15 **Quality:** Fair	**IG:** 25.1±3.8 **CG:** 26.3±4.6	BSID, Stanford-Binet	**IG:** 76.1±15.2 **CG:** 79.6±17.0	NR	After treatment, significant group difference in IQ favoring IG: ES=0.324, F(2,29)=6.96, p<0.01

Table 10. Summary of Cognitive Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Cohen et al, 2006[119] US **IG:** UCLA/Lovaas-based intervention, 21/21 **CG:** Eclectic, 21/21 **Quality:** Fair	**IG:** 30.2±5.8 **CG:** 33.2±3.7	BSID	**IG:** 61.6±16.4 **CG:** 59.4±14.7	**IG:** 87 **CG:** 73	Mean IQ increased 25 points in the treatment group and 14 points in the comparison group. The difference was significant in ANCOVA testing: F=5.21, p<0.05
		MPSMT	**IG:** 82.4±17.3 **CG:** 73.4±11.9	Mean change: **IG:** 13 **CG:** 13	No group difference, p=NS
Howard et al, 2005[111] US **IG:** UCLA/Lovaas-based intervention, 37/29 **CG1:** Intensive eclectic therapy **CG2:** Non-intensive eclectic therapy **CG2+CG3:** 41/32 **Quality:** Fair	**IG:** 30.86±5.16 **CG1:** 37.44±5.68 **CG2:** 34.56±6.53	BSID, WPPSI-R, Stanford-Binet, DAS, DAYC, PEP-R, MPSMT	Standard Scores: Composite **IG:** 70.46±11.85 **CG1:** 69.81±10.48 **CG2:** 71.62±10.47	Standard Scores: Composite **IG:** 81.32±11.14 **CG1:** 69.25±12.91 **CG2:** 68.25±9.86	Improvements in all cognitive outcomes were greater in the Lovaas group compared to comparators. All had p<0.01.
			Cognitive **IG:** 58.84±18.15 **CG1:** 53.69±13.50 **CG2:** 59.88±14.85	Cognitive: **IG:** 89.88±20.87 **CG1:** 62.13±19.63 **CG2:** 68.81±15.32	
			Non-verbal: **IG:** 80.14±11.86 **CG1:** 67.44±16.69 **CG2:** 77.69±12.33	Non-verbal: **IG:** 101.67±1 9.14 **CG1:** 73.56±24.94 **CG2:** 82.53±16.76	

* Age in months.

Abbreviations: ABA=applied behavior analysis; ACS=Autism Characteristics and Severity; ADOS=Autism Diagnostic Observation Schedule; BAS=British Abilities Scale; BSID=Bayley Scales of Infant Development; CABAS=Comprehensive Application of Behaviour Analysis; DAYC=Developmental Assessment of Young Children; DQ=developmental quotient; ESDM=Early Start Denver Model; IQ=Intelligence Quotient; LEAP=Learning Experiences and Alternative Program; MPSMT=Merrill-Palmer Scale of Mental Tests; MSEL=Mullen Scales of Early Learning; NR=not reported; PEP-R=Psychoeducational Profile; UCLA=University of California, Los Angeles; VABS=Vineland Adaptive Behavior Scale; WISC-R=Wechsler Intelligence Scale for Children-Revised; WPPSI-R=Wechsler Preschool and Primary Scale of Intelligence-Revised.

Table 11. Summary of Language Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
RCTs					
Strain et al, 2011[112] US **IG:** LEAP program with coaching and training, 28 classrooms (27 analyzed)/ 177 children **CG:** LEAP intervention manuals only, 28 classrooms (23 analyzed)/ 117 children **Quality:** Fair	**IG:** 50.1 ± 4.6 **CG:** 50.7 ± 4.2	PLS-4, total language score	**IG:** 32.8 ± 7.5 **CG:** 34.4 ± 7.2	**IG:** 51.3 ± 8.1 **CG:** 43.8 ± 7.7	Significantly greater improvement in both receptive and expressive language in the intervention group compared to the comparison. PLS-4 language score mean group differences: ES=0.92, delta=9.4, p<0.01 MSEL receptive language: ES=0.89, delta=7.3, p<0.01 MSEL expressive language: ES=0.60, delta=5.6, p<0.05
		MSEL, receptive	**IG:** 30.8 ± 7.6 **CG:** 33.4 ± 9.0	**IG:** 49.3 ± 7.9 **CG:** 40.7 ± 7.7	
		MSEL, expressive language	**IG:** 28.9 ± 7.4 **CG:** 30.3 ± 8.2	**IG:** 38.7 ± 6.4 **CG:** 35.9 ± 4.4	
Dawson et al, 2012[103,104] US **IG:** ESDM, 24/24 **CG:** Community-based interventions, 24/21 **Quality:** Good	**IG:** 23.9 ± 4.0 **CG:** 23.1 ± 3.9	MSEL, composite	**IG:** 61.0 ± 9.2 **CG:** 58.4 ± 8.6	**IG:** 78.6 ± 24.2 **CG:** 66.3 ± 15.3	After 2 years, the intervention group had significantly greater improvements in language than the comparison group. Composite score: F=4.31, MS=1264.38, p=0.044 Receptive language: F=4.14, MS=843.56, p=0.048 Expressive language: F=4.88, MS=748.07, p=0.033
		MSEL, receptive language	**IG:** 21.1 ± 4.7 **CG:** 21.2 ± 3.8	**IG:** 40 ± 16.3 **CG:** 31.5 ± 10.6	
		MSEL, expressive language	**IG:** 24.5 ± 7.2 **CG:** 26.0 ± 8.6	**IG:** 36.6 ± 13.6 **CG:** 30.0 ± 9.2	
Sallows et al, 2005[122] **IG:** Clinic directed UCLA/Lovaas-based early intensive intervention, 13/13 **CG:** Parent-directed UCLA/Lovaas-based early intensive intervention, 10/10 **Quality:** Good	**IG:** 33.23 ± 3.89 **CG:** 34.20 ± 5.06	RDLS, CELF-III	Receptive language **IG:** 38.85 ± 6.09 **CG:** 38.78 ± 6.44	**IG:** 55.85 ± 36.23 **CG:** 65.78 ± 25.81	No significant differences between groups at pre- or post-test on either measure.
			Expressive language **IG:** 47.92 ± 6.17 **CG:** 48.44 ± 6.96	**IG:** 53.38 ± 31.91 **CG:** 59.22 ± 25.13	
		VABS-Communication	**IG:** 57.46 ± 4.97 **CG:** 63.20 ± 5.58	**IG:** 73.69 ± 32.32 **CG:** 81.40 ± 24.33	
Smith et al, 2000[120] US **IG:** UCLA/Lovaas-based intervention, 15/15 **CG:** Parent training from Lovaas manual, 13/13 **Quality:** Fair	**IG:** 36.07 ± 6.00 **CG:** 35.77 ± 5.77	RDLS, comprehension	**IG:** 13.47 ± 3.60 **CG:** 13.69 ± 3.73	**IG:** 42.87 ± 22.29 **CG:** 33.0 ± 16.86	Total scores on receptive and expressive language had significantly greater improvement in the intervention group compared to the parent training group (p<0.05), although subscales did not.
		RDLS, expressive language	**IG:** 15.13 ± 0.52 **CG:** 16.31 ± 2.69	**IG:** 44.53 ± 23.48 **CG:** 36.23 ± 21.19	
		RDLS, total score	**IG:** 28.60 ± 4.07 **CG:** 30.0 ± 6.34	**IG:** 87.40 ± 46.21 **CG:** 61.33 ± 31.88	

Table 11. Summary of Language Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Nonrandomized trials					
Peters-Scheffer et al, 2013[121] Netherlands **IG:** Low intensity Lovaas-based intervention + specialized preschool, 20/20 **CG:** Eclectic preschool, 20/20 **Quality:** Good	**IG+CG:** 62.52 ± 16.96 (median)	PPVT, receptive language	**IG:** 25.00 ± 4.48 **CG:** 24.70 ± 3.21	**IG:** 34.30 ± 10.54 **CG:** 29.30 ± 7.42	Significantly greater progress in the treatment group, (repeated ANOVA): F(2,76)=3.35, p=004 ES: dividing the difference in change score by the original standard deviation of the sample, d=1.22 No significant effect on expressive language: F(1.40, 53.18)=1.50; p=0.23; d=0.40
		Schlichting Test, expressive language	**IG:** 18.35 ± 6.72 **CG:** 17.65 ± 6.64	**IG:** 34.15 ± 14.54 **CG:** 30.80 ± 15.12	
Peters-Scheffer et al, 2010[114] Netherlands **IG:** Specialized preschool+ UCLA/Lovaas-based intervention, 12/12 **CG:** Specialized preschool, 22/22 **Quality:** Fair	**IG:** 53.5 ± 5.52 **CG:** 52.95 ± 11.14	VABS-Communication	**IG:** 26.92 ± 12.12 **CG:** 25.00 ± 10.00	**IG:** 39.42 ± 15.39 **CG:** 29.92 ± 13.39	Greater improvement in IG vs. CG, p=0.02.
Hayward et al, 2009[115,116] UK **IG:** Intensive clinic-based UCLA/Lovaas-based intervention, 23/20 **CG:** Intensive parent-managed treatment, 21/19 **Quality:** Fair	**IG:** 35.7 ± 6.2 **CG:** 34.4 ± 5.7	RDLS, comprehension	**IG:** 20.0 ± 0.0 **CG:** 20.7 ± 2.8	**IG:** 26.7 ± 7.0 **CG:** 28.4 ± 9.5	Nonsignificant group differences but significant improvement on both measures between intake and followup (combined groups), p<0.01 (t-tests).
		RDLS, expressive language	**IG:** 20.2 ± 1.0 **CG:** 20.7 ± 3.3	**IG:** 26.4 ± 6.1 **CG:** 27.6 ± 7.6	
Eikeseth et al, 2002[117,118] Norway **IG:** UCLA/Lovaas-based intervention, 13/13 **CG:** Eclectic therapy, 12/12 **Quality:** Fair	**IG:** 66.31 ± 14.71 **CG:** 65 ± 10.95	RDLS, comprehension	**IG:** 49.03 ± 16.42 **CG:** 50.38 ± 15.46	**IG:** 58.47 ± 17.11 **CG:** 47.55 ± 17.25	IG showed more gains than CG on all three language measures, p<0.05.
		RDLS, expressive	**IG:** 45.12 ± 13.44 **CG:** 51.24 ± 19.24	**IG:** 67.39 ± 17.81 **CG:** 49.00 ± 18.69	
		RDLS, total	**IG:** 51.83 ± 17.42 **CG:** 60.00 ± 24.22	**IG:** 76.85 ± 26.67 **CG:** 61.58 ± 24.34	

Table 11. Summary of Language Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Cohort studies					
Boyd et al, 2013[113] US **IG:** TEACCH preschools, 85/81 **CG1:** LEAP preschools, 54/48 **CG2:** Nonmodel specific preschools, 59/56 **Quality:** Fair	**IG:** 48 ± 6.84 **CG1:** 47.52 ± 8.4 **CG2:** 48.84 ± 7.68	PLS-4, expressive communication	**IG:** 0.214 ± 0.86 **CG1:** 0.081 ± 1.045 **CG2:** -0.403 ± 0.784	**IG:** 0.441 ± 0.937 **CG1:** 0.238 ± 1.102 **CG2:** -0.317 ± 0.878	Significant baseline group difference, p<0.001. Pairwise group comparison on gain scores not significant (p=NS).
Eikeseth et al, 2012[108] Norway, Sweden **IG:** Early intervention, 35/35 **CG:** Eclectic, 24/24 **Quality:** Fair	**IG:** 3.9 ± 0.9 years **CG:** 4.4 ± 1.2 years	VABS-Communication	**IG:** 67.10 ± 14.0 **CG:** 65.5 ± 14.2	**IG:** 81.30 ± 16.90 **CG:** 63.60 ± 16.0	Greater improvement in IG vs. CG, p<0.001.
Eldevik et al, 2012[107] Norway **IG:** Preschool-based early intensive intervention, 31/31 **CG:** Usual care preschool, 12/12 **Quality:** Fair	**IG:** 42.2 ± 9.0 **CG:** 46.2 ± 12.4	VABS-Communication	**IG:** 61.90 ± 10.20 **CG:** 60.0 ± 9.60	**IG:** 70.50 ± 16.90 **CG:** 60.0 ± 14.50	Greater improvement in IG vs. CG, p<0.05. Mean ± SD change: **IG:** 8.6 ± 14.6 **CG:** 0.0 ± 12.6
Flanagan et al, 2012[109,141] Canada **IG:** Intensive behavioral intervention, 61/61 **CG:** Wait list control (matched by age), 61/61 **Quality:** Fair	**IG:** 42.93 ± 11.53 **CG:** 42.79 ± 10.51	VABS Communication	**IG:** 25.47 ± 15.81 **CG:** 25.50 ± 11.97	Mean ± SD: **IG:** 46.60 ± 29.91 **CG:** 30.33 ± 16.98 Estimated marginal mean score: **IG:** 43.45 **CG:** 29.80	Greater improvement in IG vs. CG, p=0.006. Effect size=0.56
Itzchak et al, 2011[105,106] Israel **IG:** ABA-based approach, 45/45 **CG:** Eclectic approach, 33/33 **Quality:** Fair	**IG:** 25.1 ± 3.9 **CG:** 26.0 ± 4.6	MSEL receptive language	**IG:** 34.40 ± 15.20 **CG:** 29.60 ± 14.80	**IG:** 40.10 ± 14.20 **CG:** 37.70 ± 12.80	No significant group differences on any measure.
		MSEL expressive language	**IG:** 28.80 ± 11.30 **CG:** 31.40 ± 12.50	**IG:** 35.60 ± 15.00 **CG:** 39.00 ± 14.30	
		VABS-Communication	**IG:** 67.00 ± 7.80 **CG:** 69.50 ± 10.70	**IG:** 72.90 ± 14.70 **CG:** 78.80 ± 16.20	

Table 11. Summary of Language Outcomes in Direct Provision to Child Early Intervention Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Zachor et al, 2007[110] Israel **IG:** UCLA/Lovaas-based intervention, 53/53 **CG:** Eclectic approach, 15/15 **Quality:** Fair	**IG:** 25.1 ± 3.8 **CG:** 26.3 ± 4.6	ADOS language and communication	**IG:** 13.80 ± 4.30 **CG:** 11.80 ± 4.30	**IG:** 7.20 ± 4.10 **CG:** 9.70 ± 3.00	Greater improvement in IG vs. CG, p<0.01.
Cohen et al, 2006[119] US **IG:** UCLA/Lovaas-based intervention, 21/21 **CG:** Local services, 21/19 **Quality:** Fair	**IG:** 30.2 ± 5.8 **CG:** 33.2 ± 3.7	RDLS, language comprehension	**IG:** 51.7 ± 15.2 **CG:** 52.7 ± 15.1	**IG:** 72 **CG:** 62	Significant group difference (ANCOVA): n=40, F=3.82, p<0.10 in language comprehension favoring IG. No significant group difference in expressive language: n=39, p=0.13.
		RDLS, expressive language	**IG:** 52.9 ± 14.5 **CG:** 52.8 ± 14.4	**IG:** 78 **CG:** 66	
Howard et al, 2005[111] US **IG:** UCLA/Lovaas-based intervention, 37/29 **CG1:** Intensive eclectic therapy **CG2:** Nonintensive eclectic therapy **CG2+CG3:** 41/32 **Quality:** Fair	**IG:** 30.86 ± 5.16 **CG1:** 37.44 ± 5.68 **CG2:** 34.56 ± 6.53	RDLS, receptive language	Standard Scores: **IG:** 52.16 ± 18.44 **CG1:** 45.38 ± 14.97 **CG2:** 49.00 ± 13.61	Standard Scores: **IG:** 71.31 ± 22.72 **CG1:** 49.93 ± 19.62 **CG2:** 49.21 ± 16.08	On both receptive and expressive language, the IG demonstrated significantly greater improvement, p<0.05.
		RDLS, expressive language	**IG:** 51.88 ± 12.91 **CG1:** 43.88 ± 6.69 **CG2:** 48.77 ± 11.61	**IG:** 70.46 ± 22.88 **CG1:** 47.67 ± 23.39 **CG2:** 46.79 ± 12.81	

* Age in months.

Abbreviations: ABA=Applied Behavior Analysis; ADOS=Autism Diagnostic Observation Schedule; CABAS=Comprehensive Application of Behaviour Analysis; ESDM=Early Start Denver Model; LEAP=Learning Experiences and Alternative Program; MSEL=Mullen Scales of Early Learning; NR=not reported; PLS=Preschool Language Scale; PPVT=Peabody Picture Vocabulary Test; RDLS=Reynell Developmental Language Scale; UCLA=University of California, Los Angeles; VABS=Vineland Adaptive Behavior Scale.

Table 12. Summary of Cognitive Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
RCTs					
Landa et al. 2012[124, 125] US **IG:** Assessment Evaluation and Programming System for Infants and Children (AEPS) curriculum + additional joint attention and social interaction opportunities, 25/24 **CG:** AEPS curriculum, 25/24 **Quality:** Good	**IG:** 28.6 ± 2.6 **CG:** 28.8 ± 2.8	MSEL, expressive language	**IG:** 23.92 ± 5.50 **CG:** 25.92 ± 8.12	**IG:** 34.52 ± 12.33 **CG:** 31.36 ± 12.12	No between-group differences on any measure.
		CSBS DP	**IG:** 2.29 ± 3.16 **CG:** 2.79 ± 3.62	**IG:** 8.83 ± 13.22 **CG:** 4.42 ± 4.82	
Rogers et al, 2012[129,130] US **IG:** Parent-delivered Early Start Denver model (ESDM), 49/49 **CG:** Community treatment as usual, 49/49 **Quality:** Fair	**IG:** 21.02 ± 3.51 **CG:** 20.94 ± 3.42	Mullen DQ	**IG:** 64.88 ± 17.22 **CG:** 63.08 ± 15.93	**IG:** 69.82 ± 17.9 **CG:** 67.92 ± 17.93	No significant difference between groups on DQ measures. Change from baseline, Cohen's d: Mullen DQ: **IG:** 0.44, **CG:** 0.37 Verbal DQ: **IG:** 0.56, **CG:** 0.53 Nonverbal DQ: **IG:** 0.08, **CG:** -0.01
		Mullen Verbal DQ	**IG:** 47.78 ± 22.19 **CG:** 44.45 ± 20.37	**IG:** 56.65 ± 23.65 **CG:** 54.35 ± 21.94	
		Mullen Nonverbal DQ	**IG:** 80.96 ± 16.68 **CG:** 80.73 ± 15.51	**IG:** 81.98 ± 14.82 **CG:** 80.57 ± 18.45	
Drew et al, 2002[136] UK **IG:** Parent training, 12/12 **CG:** Local/eclectic services 12/12 **Quality:** Fair	**IG:** 21.4 ± 2.7 **CG:** 23.6 ± 3.8	Griffiths Scale of Infant Development NVIQ	**IG:** 88.1 ± 11.2 **CG:** 23.6 ± 3.8	**IG:** 77.9 ± 14.8 **CG:** 66.1 ± 17.1	Significant baseline group difference, IG had a higher NVIQ than CG: $F(1, 23)=14.8$, $p<0.001$. At followup, no group differences in NVIQ.
Prospective cohort studies					
Strauss et al, 2012[126,127] Italy **IG:** Staff- and parent-mediated early intervention, 24/24 **CG:** Eclectic, 20/20 **Quality:** Good	**IG:** 55.67 ± 17.63 **CG:** 41.94 ± 13.07	Griffiths Mental Development Scale-ER GQ (Mental developmental standard score)	**IG:** 55.65 ± 20.06 **CG:** 74.29 ± 29.37	**IG:** 68.75 ± 19.58 **CG:** 76.00 ± 26.08	Change over time: IG had significant gains ($t=4.639$, $p<0.0001$). CG: $t=0.332$, $p=0.75$ Age did not influence between group difference: $F(44)=0.729$, $p=0.493$.

Table 12. Summary of Cognitive Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Reed et al, 2012[132] UK **IG:** ABA, 14/14 **CG1:** Special nursery, 21/21 **CG2:** Portage, 18/18 **CG3:** Local authority-developed parent training, 13/13 **Quality:** Fair	**IG:** 39.0 ± 6.9 **CG1:** 41.5 ± 4.0 **CG2:** 39.5 ± 6.3 **CG3:** 40.2 ± 6.3	PEP-R (Intellectual functioning)	**IG:** 55.1 ± 17.3 **CG1:** 52.2 ± 17.7 **CG2:** 54.0 ± 15.4 **CG3:** 51.7 ± 14.5	Change score **IG:** 14.5 ± 16.0 **CG1:** 10.4 ± 28.5 **CG2:** 0.6 ± 11.1 **CG3:** 3.2 ± 16.4	No between-group differences, p>0.10.

* Age in months.

Abbreviations: ABA=Applied Behavior Analysis; CG=control group; DQ=Developmental Quotient; ESDM=Early Start Denver Model; IG=intervention group; IQ=Intelligence Quotient; NVIQ=Nonverbal Intelligence Quotient; PEP-R=Psychoeducational Profile Revised.

Table 13. Summary of Language Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
RCTs					
Schreibman et al, 2013[140] US **IG:** Pivotal Response Training, 20/20 **CG:** PECS, 19/19 **Quality:** Good	**IG:** 29.5 ± 6.9 **CG:** 28.9 ± 4.2	MSEL-Expressive communication (n=38)	**IG:** 20.3 ± 3.2 **CG:** 18.5 ± 2.8	**IG:** 28.7 ± 16.5 **CG:** 23.7 ±11.2	Significant time effect for expressive communication: F=9.95, p=0.000, ES=0.216 No significant group x time interaction, F=0.551, p=0.51 Significant time effect for words produced, F=31.26, p=0.000, ES=0.486 No significant group x time interaction, F=0.313, p=0.645
		MacArthur CDI-Words produced (n=35)	**IG:** 5.3 ± 9.4 **CG:** 11.9 ± 20.5	**IG:** 129.8 ± 117.9 **CG:** 113.3 ± 108.3	
Landa et al, 2012[124,125] US **IG:** Assessment Evaluation and Programming System for Infants and Children (AEPS) curriculum + additional joint attention and social interaction opportunities, 25/24 **CG:** AEPS curriculum, 25/24 **Quality:** Good	**IG:** 28.6 ± 2.6 **CG:** 28.8 ± 2.8	MSEL	Expressive language T scores: **IG:** 23.92 ± 5.50 **CG:** 25.92 ± 8.12	Expressive language T scores: **IG:** 34.08 ± 14.59 **CG:** 31.92 ± 13.67	No significant group differences in expressive language T scores, p=0.44.
Roberts et al, 2011[131] Australia **IG:** Individualized home-based program, 34/27 **CG1:** Small group center-based program combined with parent training and support group, 33/29 **CG2:** Waitlist, 28/28 **Quality:** Fair	**IG:** 41.5 **CG1:** 43.1 **CG2:** 43.7	RDLS-Standard scores Comprehension	**IG:** 4.2 ± 9.2 **CG1:** 5.5 ± 10.6 **CG2:** 7.2 ± 15.2	**IG:** 2.6 ± 8.4 **CG1:** 10.5 ± 17.4 **CG2:** 5.7 ± 12.1	Mean group difference, 95% CI **IG vs. CG1:** 7.3 (0.7 to 13.9), p=0.03 **IG vs. CG2:** 1.8 (-4.9 to 8.4), p=0.60 **CG1 vs. CG2:** 5.5 (-1.2 to 12.2), p=0.10 3-group comparison: p=0.08
		Expression Standard scores	**IG:** 3.4 ± 8.3 **CG1:** 8.2 ± 16.6 **CG2:** 6.0 ± 10.9	**IG:** 2.8 ± 7.5 **CG1:** 7.0 ± 15.1 **CG2:** 4.4 ± 8.7	Mean group difference, 95% CI **IG vs. CG1:** 3.0 (-2.9 to 9.0), p=0.31 **IG vs. CG2:** 0.9 (-4.9 to 6.8), p=0.75 **CG1 vs. CG2:** 2.1 (-3.8 to 8.0), p=0.48 3-group comparison: p=0.58 Significant greater improvement in Reynell comprehension standard score for CG1 compared with IG. All other group differences were not significantly different for both comprehension and expression scores.

Table 13. Summary of Language Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Aldred et al, 2012[137,138] UK **IG:** Parent training in social communication intervention plus community intervention, 14/14 **CG:** Eclectic intervention, 14/14 **Quality:** Good	**IG:** 51.4 ± 11.8 **CG:** 50.9 ± 16.3	MacArthur CDI-Language comprehension	**IG:** 71.7 ± 2383 **CG:** 95.4 ± 426	**IG:** 222.7 ± 40,431 **CG:** 146.8 ± 11,426	No significant group difference in language comprehension: F=2.93, p=0.1. Significant group difference in expressive language: F=18.5, p<0.001.
		Expressive language	**IG:** 28 ± 467 **CG:** 25.6 ± 683	**IG:** 199.4 ± 25,606 **CG:** 33.1 ± 683	
Casenhiser et al, 2013[128] Canada **IG:** MEHRIT (developmental individualized relationship-based intervention), 25/25 **CG:** Eclectic community-based treatment, 26/26 **Quality:** Fair	**IG:** 42.5 ± 8.8 **CG:** 46.4 ± 8.3	PLS-4, CASL	Developmental quotient (DQ): **IG:** 0.64 ± 0.32 **CG:** 0.54 ± 0.26	DQ: **IG:** 0.72 ± 0.39 **CG:** 0.64 ± 0.32	Both groups significantly different from baseline scores. **IG:** p=0.038, d=0.451 **CG:** p<0.001, d=0.915 No significant group difference in DQ scores after treatment: F(1,48)=1.589, p=0.214, η^2=0.022
Oosterling et al, 2010[133] **IG:** Nonintensive parent training+specialized preschool, 40/36 **CG:** Eclectic preschool, 35/31 **Quality:** Fair	**IG:** 35.2 ± 5.5 **CG:** 33.3 ± 6.4	MacArthur CDI Words understood	**IG:** 177.9 ± 122.5 **CG:** 181.5 ± 121.4	Change from baseline **IG:** 62.0 ± 75.0 **CG:** 35.2 ± 66.1	No significant group difference in words understood: F=2.12, p=NS No significant group difference in words said: F=2.92, p=NS Significant group difference in gestures produced: F=3.91, p=NS
		Words said	**IG:** 106.8 ± 122.2 **CG:** 101.7 ± 109.7	**IG:** 75.5 ± 78.8 **CG:** 56.1 ± 97.2	
		Gestures produced	**IG:** 29.1 ± 13.7 **CG:** 30.1 ± 13.6	**IG:** 6.7 ± 10.2 **CG:** 6.3 ± 9.0	
Rogers et al, 2012[129,130] US **IG:** Parent-delivered Early Start Denver Mode (ESDM), 49/49 **CG:** Eclectic community treatment, 49/49 **Quality:** Fair	**IG:** 21.02 ± 3.51 **CG:** 20.94 ± 3.42	MacArthur CDI Part I: Phrases Understood	**IG:** 8.22 ± 7.02 **CG**: 9.38 ± 7.95	**IG:** 12.73 ± 9.11 **CG:** 14.77 ± 8.14	Baseline vs. Visit 2, Cohen's d: Phrases Understood: **IG:** 0.62, **CG:** 0.87 Vocabulary Comprehension: **IG:** 0.66, **CG:** 0.84 Vocabulary Production: **IG:** 0.69, **CG:** 0.57 Total Gestures: **IG:** 0.83, **CG:** 1.02 Significant group difference in number of intervention hours (p<0.05) and Intervention hours had a significant interaction effect on MacArthur CDI Vocabulary production (p=0.005)
		MacArthur CDI Part I: Vocabulary Comprehension	**IG:** 64.53 ± 65.73 **CG:** 70.31 ± 78.34	**IG:** 106.51 ± 96.81 **CG:** 125.72 ± 106.39	
		MacArthur CDI Part I: Vocabulary Production	**IG:** 12.24 ± 35.6 **CG:** 12.44 ± 39.72	**IG:** 42.27 ± 61.99 **CG:** 38.87 ± 73.71	
		MacArthur CDI Part II: Total Gestures	**IG:** 19.89 ± 10.12 **CG:** 20.33 ± 11.15	**IG:** 28.02 ± 12.62 **CG:** 29.79 ± 13.51	

Table 13. Summary of Language Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Carter et al, 2011[134] US **IG:** More than Words, 32/29 **CG:** Eclectic, 30/26 **Quality:** Fair	**IG:** 21.11 ± 2.71 **CG:** 21.51 ± 2.82	Mullen Expressive Language Age	**IG:** 8.22 ± 6.01 **CG:** 7.33 ± 3.71	**IG:** 16.20 ± 7.23 **CG:** 16.68 ± 7.88	No significant between group differences. No significant group difference in PIA-CV nonverbal communication residualized Gain scores: ES=-0.19 (95% CI, -0.81 to 0.43). No significant group difference in ESCS initiating joint attention residualized Gain scores: ES=0.12 (95% CI, -0.46 to 0.70). No significant group difference in PCFP frequency of intentional communication residualized Gain scores: ES=0.15 (95% CI, -0.57 to 0.88).
		Mullen Receptive Language Age	**IG:** 8.41 ± 5.42 **CG:** 8.17 ± 4.44	**IG:** 15.52 ± 6.93 **CG:** 17.48 ± 8.33	
		PIA-CV nonverbal communication	**IG:** 2.30 ± 0.64 **CG:** 2.28 ± 0.73	**IG:** 2.89 ± 0.67 **CG:** 2.92 ± 0.65	
		ESCS initiating joint attention	**IG:** 5.90 ± 5.41 **CG:** 5.59 ± 6.14	**IG:** 10.33 ± 9.82 **CG:** 8.68 ± 9.26	
		PCFP weighted frequency of intentional communication	**IG:** 5.55 ± 6.29 **CG:** 8.20 ± 12.63	**IG:** 18.91 ± 20.50 **CG:** 20.75 ± 21.14	
Green et al, 2010[135] UK **IG:** Preschool autism communication intervention (PACT), 77/74 **CG:** Eclectic treatment, 75/72 **Quality:** Fair	**IG:** 45 **CG:** 45	PLS-Receptive raw scores	**IG:** 15.6± 9.8 **CG:** 15.0±9.7	**IG:** 21.5±13.0 **CG:** 20.3 ±12.8	Change from baseline, mean (SD): **IG:** 6.0 (6.7), **CG:** 5.3 (5.9) Change from baseline, mean (SD): **IG:** 5.1 (5.6), **CG:** 4.9 (5.2) Change from baseline, mean (SD): **IG:** 74.2 (66.9), **CG:** 47.0 (68.2) Change from baseline, mean (SD): **IG:** 78.5 (89.3), **CG:** 51.8 (73.2) Significant group differences, mean (95% CI): 30.28 (6·90 to 53·68) for the MacArthur CDI receptive scores and 21.37 (-6·42 to 49·16) for the MacArthur CDI expressive scores
		Expressive raw scores	**IG:** 15.0±8.1 **CG:** 15.1±7.9	**IG:** 20.0 ±11.2 **CG:** 20.0±11.3	
		Parent-rated MacArthur CDI Receptive raw score	**IG:** 159.5±114.4 **CG:** 162.0± 122.4	**IG:** 233.7±129.6 **CG:** 209.0±131.3	
		MacArthur CDI Expressive raw score	**IG:** 93.5 ± 114.8 **CG:** 111.1 ± 128.6	**IG:** 171.9 ± 150.7 **CG:** 163.8 ± 144.3	
Drew et al, 2002[136] UK **IG:** Parent training, 12/12 **CG:** Eclectic services, 12/12 **Quality:** Fair	**IG:** 21.4 ± 2.7 **CG:** 23.6 ± 3.8	MacArthur CDI-Words understood	**IG:** 52 ± 60.5 **CG:** 53.0 ± 63.7	**IG:** 176.1 ± 121.9 **CG:** 100.3 ± 80.2	Words and gestures produced were not significantly different between the 2 groups (p >0.05). Significantly more children in the IG moved from being nonverbal to having single word or phrase speech (n=7) than in the CG (n=2; 1 child went from single words to <5 words from initial to followup assessment) (Fisher exact test, p<0.05).
		Words said	**IG:** 6.8 ± 20.9 **CG:** 6.6 ± 13.7	**IG:** 96.6 ± 118.8 **CG:** 44.0 ± 50.2	
		Total gestures produced	**IG:** 20.9 ± 7.0 **CG:** 20.9 ± 14.4	**IG:** 38.6 ± 12.5 **CG:** 29.1 ± 18.4	
		ADI-R overall language rating	**IG:** Nonverbal (<5 words): 11 Single words: 1 Phrase speech: 0 **CG:** Nonverbal (<5 words): 11 Single words: 1 Phrase speech: 0	**IG:** Nonverbal (<5 words): 4 Single words: 5 Phrase speech: 3 **CG:** Nonverbal (<5 words): 9 Single words: 3 Phrase speech: 0	

Table 13. Summary of Language Outcomes in Parent Training Studies

Author, Year, Country Groups, n Enrollment/ n Final	Age at Intake,* Mean±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Prospective cohort studies					
Strauss et al, 2012[126,127] Italy **IG:** Staff and parent mediated early intervention, 24/24 **CG:** Eclectic, 20/20 **Quality:** Good	**IG:** 55.67 ± 17.63 **CG:** 41.94 ± 13.07	MacArthur CDI - Comprehension	**IG:** 53.83 ± 28.81 **CG:** 47.17 ± 27.80	**IG:** 70.33 ± 27.04 **CG:** 61.33 ± 32.37	Comprehension: Change over time significant for both groups, p≤0.012. Children in IG outperformed children in CG at T2 on early language comprehension: t(23)=6.460, p≤0.001 vs. t(19)=3.885, p≤0.05. Production: Change over time significant for IG (t=3.41, p=0.003) but not for CG (t=1.69, p=0.151) No age effect influencing between-group differences in CDI language comprehension: F(44)=1.492, p=0.249) or CDI language production: F(44)=1.553, p=0.233.
		Production	**IG:** 35.29 ± 35.97 **CG:** 19.17 ± 28.12	**IG:** 51.81 ± 35.23 **CG:** 33.17 ± 42.27	
Keen et al, 2010[139] Australia **IG:** Professional parent intervention, 17 families/NR **CG:** Self-directed video based parent intervention, 22 families/NR **Quality:** Good	**IG:** 36.38 ± 7.54 **CG:** 35.71 ± 6.92	CSBS DP Caregiver	**IG:** 69.10 ± 24.16 **CG:** 62.82 ± 28.01	**IG:** 91.67 ± 28.80 **CG:** 80.83 ± 27.68	Greater improvement in IG vs. CG on parent-rated caregiver measure but not observer-rated behavior sample.
		CSBS DP Behavior Sample	**IG:** 56.36 ± 31.84 **CG:** 55.57 ± 38.24	**IG:** 75.84 ± 39.73 **CG:** 73.57 ± 48.92	

* Age in months.

Abbreviations: ADI-R=Autism Diagnostic Interview Revised; CASL=Comprehensive Assessment of Spoken Language; CSBS DP=Communication and Symbolic Behavior Scales Developmental Profile; MacArthur CDI=MacArthur Communication Development Inventory; CG=control group; CI=confidence interval; ES=effect size; IG=intervention group; MSEL=Mullen Scales of Early Learning; PACT=Preschool Autism Communication Intervention; PECS=Picture Exchange Communication System; PLS-4=Preschool Language Scale; RDLS=Reynell Developmental Language Scales.

Table 14. Key Outcomes of Randomized, Controlled Trials of Play/Interaction-Based Interventions

Author, Year, Country Groups, n Enrollment/n Final Study Quality	Age,* Mean ± SD IQ, Mean ± SD	Key Outcomes
Joint attention studies		
Kasari et al, 2014[142] US **IG:** Caregiver-mediated joint attention/symbolic play intervention, 60/51 **CG:** Caregiver education, 52/44 **Quality:** Fair	**IG:** 41.9 ± 10 **CG:** 42.8 ± 10.21 **IG:** 23.6 ± 11.6 **CG:** 26.3 ± 11.8	• 66% of participants were non-white and from families with low SES status • Joint engagement improved in both groups over time with significantly greater improvement in IG vs. CG, p=0.003 • Time engaged increased to over half of interaction period in IG, rate of difference between groups=44.7%. Effect size for difference in joint engagement=0.21. Differences in joint engagement maintained at followup 3-months post-intervention for IG, p=0.02 • Initiating joint attention improved in both groups with greater improvement in IG, p=0.05, effect size=0.14. Increases in both groups maintained at followup • No significant improvements in functional play types in either group
Goods et al, 2013[155] US **IG:** Joint attention intervention, 8/6 **CG:** Control, 7/5 **Quality:** Good	**IG:** 48.73 ± 11.68 **CG:** 54.68 ± 10.25 **IG:** 37.70 ± 15.21 **CG:** 26.67 ± 10.12	• IG demonstrated more spontaneous play types, spent less time unengaged in classroom, and initiated more requesting gestures than CG (effect sizes 0.81, 1.63, and 1.51 respectively, p values ≤0.05) • No significant group differences on the Early Social Communication Scales measures of joint attention
Kasari et al, 2010[154] US **IG:** Immediate joint attention intervention, 19/19 **CG:** Waitlist control, 19/19 **Quality:** Good	**IG:** 30.35 ± 0.93 **CG:** 31.31 ± 0.90 **IG:** 64.80 ± 5.35 **CG:** 59.81 ± 3.14	• Children in IG exhibited significantly less object-focused play, more responsiveness to joint attention, more functional play types, and greater joint engagement than CG at initial followup (p<0.05); gains in joint engagement, responsiveness to joint attention, and types of functional play were maintained at 1-year followup of IG • Groups did not differ on other/unengaged play time at followup
Wong 2013[143] US **IG:** Joint attention-symbolic play interventions, 14/14 **CG:** Symbolic play-joint attention intervention, 10/10 **G3:** Waitlist, 9/9 **Quality:** Fair	**IG:** 56.2 ± 10.4 **CG:** 54.5 ± 5.1 **G3:** 59.7 ± 10.6 NR	• Increased time in joint engaged state for IG and CG vs. G3 (effect size=0.63) • For IG+CG, joint engagement time, joint attention responses/minute, joint attention initiations/minute, symbolic play acts/minute increased significantly from baseline to post-intervention (effect sizes of 0.41, 0.43, 0.21, and 0.51 respectively) • Increases in joint attention responses from baseline for IG+CG as measured on the Early Social Communication Scales; no significant increases in functional play level or structured play • No significant modifiers identified
Warreyn et al, 2013[144] Belgium **IG:** Joint attention/imitation intervention, 18/18 **CG:** Treatment as usual, 18/18 **Quality:** Fair	**IG:** 5.7 ± 0.6 years **CG:** 5.7 ± 0.7 years **IG:** 78.9 ± 15.5 **CG:** 76.9 ± 16.8	• Total joint attention scores more improved for IG vs. CG (p<0.01); gaze following, initiating requests also significantly improved for IG vs. CG (p values <0.05) • IG increased number of elicited joint attention acts by 1.88, number spontaneous declarative joint attention actions by 0.83, and number correct imitations by 7.01 from baseline in both groups from baseline to followup (p<0.05) • Verbal IQ significantly correlated with growth in imitation for IG (p<0.05); age, mental age, full scale IQ baseline imitation and joint attention skills, performance IQ were not significant modifiers of outcomes

Table 14. Key Outcomes of Randomized, Controlled Trials of Play/Interaction-Based Interventions

Author, Year, Country Groups, n Enrollment/n Final Study Quality	Age,* Mean ± SD IQ, Mean ± SD	Key Outcomes
Schertz et al, 2013[153] US **IG:** Joint attention-focused parent training, 11/11 **CG:** Treatment as usual, 12/12 **Quality:** Fair	**IG:** 24.6 ± 4.0 **CG:** 27.5 ± 3.4 NR	Scores on responding to joint attention significantly improved for IG vs. CG at 4-week post-intervention followup (effect size for differences=1.39), as were scores on focusing on faces (effect size=1.24); effects sizes at 8-week followup were 1.18 (responding to joint attention) and 0.84 (faces)
Lawton et al, 2012[148] US **IG:** Immediate joint attention intervention, 9/9 **CG:** Delayed treatment, 7/7 **Quality:** Fair	**IG:** 46.0 ± 5.00 **CG:** 43.01 ± 6.00 **IG:** 30.3 ± 5.01 **CG:** 33.8 ± 8.74	• Joint attention intervention delivered by preschool teachers • In classroom observations, IG demonstrated greater initiations of joint attention vs. CG (effect size=1.85, p<0.005) and used more pointing and showing gestures (effect sizes 2.02 and 1.85 respectively); no differences in looking or giving • Total joint attention scores on the Early Social Communication Scales did not differ between groups • On intervention exit play observations, no group differences in any joint attention skills • IG demonstrated less object engagement (effect size=1.41) and more supported engagement (effect size=1.24) compared with CG
Kaale et al, 2012[146] Norway **IG:** Joint attention intervention, 34/34 **CG:** Control, 27/27 **Quality:** Fair	**IG:** 47.6 ± 8.30 **CG:** 50.3 ± 8.3 **IG:** 53.3 ± 19.2 **CG:** 59.9 ± 19.7	• Joint attention intervention delivered by preschool teachers • IG demonstrated more frequent joint attention skills in play with teachers vs. CG, with IG nearly 5 times more likely to demonstrate initiation of joint attention vs. CG (effect size=0.44); duration of joint engagement with teachers did not differ between groups • IG spent longer time in jointly engaged play with mothers vs. CG post-intervention (effect size=0.67); frequency of joint attention skills with mothers did not differ between groups
Kasari et al, 2012[149-152] US **IG:** Joint attention intervention, 20/20 **CG:** Symbolic play intervention, 16/16 **G3:** Control, 16/16 **Quality:** Fair	**IG:** 43.05 ± 6.863 **CG:** 41.41 ± 6.491 **G3:** 41.31 ± 4.542 NR	• Joint attention interventions delivered by interventionists; children in the intervention groups showed greater growth in expressive language, initiation of joint attention, and duration of child-initiated joint attention than did control group children (p<0.01, <0.05); receptive language growth not significantly affected by intervention • Amount of intervention services received post-intervention was not related to growth in skills at followup 12 months after the ~6 week intervention, except for child-initiated joint attention: children receiving fewer hours of additional services showed greater growth in child-initiated joint attention • Quality of joint attention (shared positive affect, shared positive affect with utterances) improved in IG and CG at 6 and 12 month followups • At followup of 40/58 participants 5 years post-intervention, 32/40 had passing scores on the Expressive Vocabulary Test of spoken language; only baseline play level predicted ability to use spoken language. • Younger age at baseline, initiation of joint attention, and play level were predictors of spoken language ability at 5-year followup • Greater functional play types at baseline predicted better overall cognitive ability at 5-year followup

Table 14. Key Outcomes of Randomized, Controlled Trials of Play/Interaction-Based Interventions

Author, Year, Country Groups, n Enrollment/n Final Study Quality	Age,* Mean ± SD IQ, Mean ± SD	Key Outcomes
Imitation studies		
Ingersoll 2010[156,157] US **IG:** Reciprocal imitation training, 15/14 **CG:** Control, 14/13 **Quality:** Good	**IG:** 41.36 ± 4.30 **CG:** 37.20 ± 7.36 NR	• Pilot evaluation of a reciprocal imitation training program • IG made greater gains in spontaneous and prompted imitation, object imitation, gesture imitation, initiation of joint attention, and on the Social-Emotional Scale than CG (p values ≤0.05) • Number of spontaneous play actions associated with gains in spontaneous imitation and gesture imitation (p<0.05) • Changes in imitation skills not associated with social functioning changes in mediation analysis
Parent-child communication studies		
Siller et al, 2013[147] US **IG:** Parental responsiveness intervention, 36/31 **CG:** Control, 34/31 **Quality:** Fair	**IG:** 58.3 ± 12.7 **CG:** 55.9 ± 11.9 NR	• Intervention focused on increasing parents' responsiveness to child communication • Mothers of children in IG demonstrated greater synchronization with child communication vs. CG (p<0.05, effect size=0.08) • No significant effects of intervention on expressive language • Mothers rated as more insightful at baseline had greater gains in synchronization
Venker et al, 2011[145] US **IG:** Parental responsiveness intervention, 7/7 **CG:** Delayed treatment, 7/7 **Quality:** Fair	**IG+CG:** 41.14 ± 10.40 NR	• Intervention targeting parents' verbal responsive and engagement with child play • Both groups increased prompted communication acts from baseline to followup; in between group comparisons, IG had greater increases vs. CG (p<0.03) • Number of children increasing spontaneous communication acts did not differ between groups

* Age in months.

Abbreviations: BASC=Behavioral Assessment System for Children; CG=control group; DQ=Developmental Quotient; IG=intervention group; IQ=intelligence quotient; N=number; NR=not reported; SD=standard deviation.

Table 15. Summary of Meta-Analyses of Early Intervention Approaches

Author, Year	Study Type as Defined in Review (n) Total Participants/Group (n)	Mean Participant Age (Months)	Treatment Intensity, Hours/Week Treatment Duration, Mean Months (Range)	Effect Sizes (95% CI)
Reichow 2012[167]	RCT: 1 Controlled trial: 4 Early intervention: 116 Comparison: 87	30.2 to 42.5	>24 hours/week 26.3 months (14 to 36)	IQ: 0.76 (0.40 to 1.11) Expressive language: 0.50 (0.05 to 0.95) Receptive language: 0.57 (0.20 to 0.94) Vineland adaptive behavior: 0.69 (0.38 to 1.01) Vineland communication: 0.74 (0.30 to 1.18) Vineland socialization: 0.42 (0.11 to 0.73) Vineland daily living: 0.55 (0.24 to 0.87)
Virues-Ortega 2010[169]	Total studies (type not defined): 22 Early intervention: 323 Comparison: 180	22.6 to 66.3	12 to 45 hours/week 4 to 34 months	IQ: 1.19 (0.91 to 1.47) Expressive language: 1.47 (0.85 to 2.08) Receptive language: 1.48 (0.96 to 1.97) General language: 1.07 (0.34 to 1.79) Vineland adaptive behavior: 1.09 (0.70 to 1.47) Vineland socialization: 0.95 (0.53 to 1.37) Vineland communication: 1.45 (1.02 to 1.88) Vineland daily living: 0.62 (0.30 to 0.93) Vineland motor skills: 0.71 (0.19 to 1.22)
Peters-Scheffer 2011[168]	RCT: 1 Pre-test/post-test with control: 10 Early intervention: 168 Comparison: 144	33.65 to 65.68	12.5 to 38.6 hours/week 10 to 24+ months	IQ: 2.00 Nonverbal IQ: 0.98 Expressive language: 1.10 Receptive language: 2.91 Vineland adaptive behavior: 0.91 Vineland communication: 1.32 Vineland daily living: 0.68 Vineland socialization: 1.49

Note: Of the 30 early intensive behavioral and developmental studies described in this current review, four[117-119,122,123] were included in three of the prior reviews[167-169,171] summarized in this report, and two[19,111] were included in four[167-169,171] of the reviews. One additional study[107] was included in one prior review.[169]

Abbreviations: CI=confidence interval; IQ=intelligence quotient; RCT=randomized, controlled trial.

Table 16. Summary of Evidence

Key Question	Number of Studies Overall Quality	Limitations	Consistency	Applicability	Summary of Findings
1. Outcomes of screening	No studies identified	NA	NA	NA	No studies directly link screening to health outcomes
2. Performance characteristics	18[39,58,59,61-79] Overall quality: Fair	-Few studies followed participants who screened negative to identify false negatives -All studies demonstrated significant loss to followup across multiple steps from screening to diagnosis	Consistent	-Fair generalizability -Population-based screening (well-child visits) -Studies of M-CHAT conducted in US, Sweden, Spain, Japan, Australia	-PPV for ASD in the 2 large studies of the most common screening process (M-CHAT-R and M-CHAT-R/F) was around 50% -Few children receiving diagnostic evaluations were typically developing as most had some non-ASD developmental disorder
2a. Risk factors that modify screening performance	No studies identified	NA	NA	NA	NA
2b. Age as modifier of screening performance	1[70]	Screen negatives not systematically followed	NA	-Fair generalizability -Population-based screening in US	PPV was higher for children age >24 months. Sample excluded many children with ASD, limiting the ability of this analysis to comment on true population estimates.
2c. Other characteristics that modify screening performance	No studies identified	NA	NA	NA	NA
3. Harms of screening	No studies identified	NA	NA	NA	NA
4. Interventions for young children with ASD	42 studies assessing interventions aimed at younger children; 5 systematic reviews Overall quality: Fair	-Multiple outcome measures -Effects of potential modifiers unclear -Participants often received multiple interventions, thus disentangling effects is difficult	Consistent	-Good generalizability -Overall studies were conducted in populations and settings similar to U.S. health care context. Many interventions broadly accessible in the United States	Children in early intensive behavioral and developmental intervention made gains in cognitive and language skills (significant gains for treatment over control groups in 11 of 20 studies assessing cognitive outcomes; 13 of 28 studies assessing language; and 9 of 10 studies addressing joint attention outcomes)
4a. Effects of intervention timing on treatment outcomes	4[107,109,129,138] Overall quality: Fair	-Few studies had specific modifier analyses -Outcome measures varied across studies	Consistent	-Good generalizability -Overall studies were conducted in populations and settings similar to U.S. health care context. Many interventions broadly accessible in the United States	-Younger age at intervention was associated with greater gains. Examples include: all children in 1 study with IQ gains of at least 30 were age <6, with a mean of 3.5; in another study, age was an independent predictor in multivariate analyses of shifts in IQ at $p<0.001$; in a third, baseline age (younger toddlers) was associated with a significant shift in developmental quotient, and in a fourth,

Table 16. Summary of Evidence

Key Question	Number of Studies Overall Quality	Limitations	Consistency	Applicability	Summary of Findings
					age positively correlated with changes in adaptive behavior on the ABC scale. Nonetheless, studies need to be replicated as different interventions and outcome scales were used, and some other studies found no difference by age, particularly those with tight age bands in the preschool age group.
4b. Effect of ASD severity on intervention outcomes	3[105-107] Overall quality: Fair	-Few studies had specific modifier analyses -Outcome measures varied across studies	Consistent	-Good generalizability -Overall studies were conducted in populations and settings similar to U.S. health care context. Many interventions broadly accessible in the United States	-Having a diagnosis on the ASD spectrum other than autism was associated with greater gains (e.g. Asperger Syndrome or PDD-NOS). Other data are inconsistent on whether symptom severity is associated with outcome achievement.
5. Harms of intervention	No studies of interventions aimed at younger children[a]	No studies of behavioral interventions reported harms	NA	NA	NA

[a] Studies of medical interventions—typically including older children—reported significant harms.

Abbreviation: ASD=autism spectrum disorder.

Appendix A. Search Strategies

Table A1. PubMed Search Strategies

Search terms		Search results
#1	autistic[tiab] OR autism[tiab] OR autistic disorder[mh] OR asperger syndrome[mh] OR child development disorders, pervasive[mh:noexp] OR asperger[tiab] OR asperger's[tiab] OR aspergers[tiab] OR pervasive development[tiab] OR pervasive developmental[tiab]	24402
#2	mass screening[mh] OR screening[tiab] OR screened[tiab] OR screen[tiab] OR screener[tiab] OR screeners[tiab] OR early diagnosis[mh] OR identify[tiab] OR identification[tiab]	1219531
#3	#1 AND #2 AND eng[la] AND humans[mh]	2325
#4	newspaper article[pt] OR letter[pt] OR comment[pt] OR case reports[pt] OR review[pt] OR practice guideline[pt] OR news[pt] OR editorial[pt] OR historical article[pt] OR meta-analysis[pt] OR legal cases[pt] OR jsubsetk	4819874
#5	#3 NOT #4 AND 2000:2014[dp]	1511

Key: [mh] Medical Subject Heading; [tiab] title/abstract word; [pt] publication type; [dp] publication date; jsubsetk consumer health literature

Table A2. PsycINFO Search Strategies (ProQuest Interface)

Search terms		Search results
#1	SU.EXACT.EXPLODE("pervasive developmental disorders" or "aspergers syndrome" or "autism")	23790
#2	SU.EXACT.EXPLODE(screening) or SU.EXACT.EXPLODE(screening tests) or screening or screened or screen or screener or screeners or identify or identification	232628
#3	#1 and #2 and DTYPE(journal article) and (ME(empirical study) or ME(field study) or ME(followup study) or ME(longitudinal study) or ME(prospective study) or ME(qualitative study) or ME(quantitative study) or ME(retrospective study) or ME(treatment outcome/clinical trial)) and LA(English), limited to peer-reviewed journals and human population, limited to publication date 2000 to present	1267

Key: SU.EXACT.EXPLODE subject descriptor; DTYPE publication type; ME methodology; LA language

Table A3. ERIC Search Strategies (ProQuest Interface)

Search terms		Search results
#1	("pervasive developmental disorders") or autism or autistic or ("asperger syndrome")	11548
#2	screening or screened or screen or screener or screeners or identify or identification	92626
#3	#1 and #2, limited to peer reviewed journals, scholarly journals, English only, publication date 2000 to present	1089

Table A4. CINAHL Search Strategies (EBSCO Interface)

Search terms		Search results
#1	(MH "Child Development Disorders, Pervasive") OR (MH "Asperger Syndrome") OR (MH "Autistic Disorder")	8841
#2	(MH “Early Diagnosis”) OR (MH “Health Screening”) or screening or screened or screen or screener or screeners or identify or identification	147216
#3	#1 AND #2 AND PT systematic review	23
#4	#1 AND #2 AND PT review	49
#5	#1 AND #2 AND PT case study	33
#6	#1 AND #2 NOT (#3 OR #4 OR #5), limited to English, research articles, peer-reviewed journals, and publication date 2000 to present, excluding MEDLINE records	33

Key: MH subject term; PT publication type

Screening for ASD Systematic Evidence Review
Abstract Review Form

First Author, Year: _____ Reference ID #: _____ Initials: ___ ___ ___

1. Does the study address screening measures for ASD?	Yes	No
2. Is the study original research (includes systematic review or meta-analysis)?	Yes	No
3. Does the study include individuals in the target population (children screened for ASD between 12-36 months of age)?	Yes	No
Record age range of participants:		
Record the mean age and standard deviation of participants:		
Record the total N participants:		
If excluded, retain for review of references or contextual questions?	Yes	No

Screening for ASD Systematic Evidence Review
Full Text Review Form

First Author, Year: _____ Reference ID #: _____ Initials: ___ ___ ___

1. Does the study address screening measures for ASD?	Yes	No
2. Is the study original research (includes systematic review or meta-analysis)?	Yes	No
3. Does the study include individuals in the target population (children screened for ASD between 12-36 months of age)? a. Record age range (mean+SD and/or range)	Yes	No
4. Does the study provide data related to at least one of the following? • Performance characteristics of tests to screen for ASD in children between 12-36 months of age (provides sensitivity, specificity, PPV, NPV or data to calculate) • Harms or long-term clinical outcomes of screening for children between 12-36 months of age/families of such children (e.g., behavioral changes, distress, misclassification, other unintended consequences)	Yes	No
5. Record setting in which screening occurs (e.g. primary care, early intervention preschool)		
6. Was the study conducted in one of the following countries: Andorra, Argentina, Australia, Austria, Bahrain, Barbados, Belgium, Brunei Darussalam, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong (China), Hungary, Iceland, Ireland, Israel, Italy, Japan, Republic of Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Qatar, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States	Yes	No
7. Does the study include an exposure/screening group and comparison group? (Note: This question for data collection only; does not affect inclusion)	Yes	No
8. If excluded, retain this paper for contextual questions or review of references (performance characteristics of diagnostic tools, stability of diagnosis, resources available for screening, treatment resources available)?	Yes	No

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Chlebowski et al, 2013[1]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:** • Excluded if received an ASD diagnosis before being screened with the M-CHAT • Severe sensory or motor disability (e.g., blindness or deafness) • child's caregivers not fluent in English or Spanish • Includes children from early intervention and pediatric referrals Note: See related publications Robins 2008,[2] Kleinman 2008,[3] Pandey 2008,[4] Robins 2001[5]	**Population characteristics** **Population description**: Screening of 18 989 toddlers at pediatric well-child visits (18-24 month) in 2 US geographic regions **Race/ethnicity, n (%):** White, non-Hispanic: 6184 (68.4) Nonwhite: 1186 (13.1) **Age at screening, mean ± SD:** 20.4 ± 3.1 months **Sex, n (%):** Female: 9388 (49.4) Male: 9601 (50.6) **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Mental age:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context** SES: NR Insurance status: NR Parental education: NR Marital status: NR Primary language spoken in home: NR Limited English proficiency: NR	**Screening** **Instrument/method**: M-CHAT+ phone followup (M-CHAT/F) **Mode of administration:** Parent-rated, M-CHAT-F by phone **Setting**: Primary care **Administrator:** Research assistants **Scorer:** Research office **Study definition of positive screen**: Screening positive on 2 of 6 critical items or on 3 of 23 items overall on both the M-CHAT and M-CHAT/F **Total number available to screen**: 18989 **Total number screened**: With M-CHAT: 18989 With M-CHAT/F: 1295 M-CHAT: Number screening positive: 1737 Number screening negative: 17252 M-CHAT/F: Number screening positive: 272 Number screening negative: 1023 **Percent completing screening followup/diagnosis:** 1.1% (207/18989) **Diagnosis** **Diagnostic process:** ADOS, ADI-R, Mullen Scales of early Learning, the Vineland Adaptive Behavior Scales and CARS, DSM-IV; included those who screened positive on M-CHAT/F (n=171) and those flagged by pediatrician despite screen negative (n=36) **Who diagnosed**: Licensed clinical psychologist or developmental pediatrician and a psychology doctoral student	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD**: NR **Language development, mean ± SD**: NR **Adaptive behavior, mean ± SD:** NR **Family context** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors**: NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD**: NR **Language development, mean ± SD:** NR **Family context:** NR	**Performance characteristics (%)** **M-CHAT:** PPV: 6% **M-CHAT + M-CHAT/F:** PPV: 54% **Other outcomes** **Intermediate/Process outcomes:** NR **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Chlebowski et al, 2013[1]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Diagnoses, n: ASD: Screening Positive: 92 (autistic disorder=44, PDD-NOS=48) Screening Negative: 1 Non-ASD diagnoses, n: Developmental Delay: Screening Positive: 37 Screening Negative: 11 Developmental language disorder: Screening Positive: 18 Screening Negative: 6 Other DSM-IV diagnoses: Screening Positive: 5 Screening Negative: 1 Developmental concerns: Screening Positive: 15 Screening Negative: 9 Typical development: Screening Positive: 4 Screening Negative: 3 **Age at diagnosis, mean ± SD:** 25.75 ± 4.51, n=207		

Study Description Author/Year: Kleinman et al, 2008[3]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:** • Excluded if received an ASD diagnosis before being screened with the M-CHAT • Severe sensory or motor disability (eg, blindness or deafness) • child's caregivers not fluent in English or	**Population characteristics** **Population description:** Initial screen: Screening of toddlers at pediatric well-child visits or during early intervention intake or after referral from developmental pediatrician or psychologist Re-screening: children who passed initial screening or followup interview; also included participants from Robins et al, 2001 **Race/ethnicity, n (%):** NR **Age at screening, mean months ± SD:** 21.01 ± 3.37	**Screening** **Instrument/method:** M-CHAT+ phone followup (M-CHAT/F) at 2 time points **Mode of administration:** Parent-rated, M-CHAT-F by phone **Setting:** Primary care **Administrator:** Research assistants **Scorer:** Research office **Study definition of positive screen:** Screening positive on 2 of 6 critical items or on 3 of 23 items overall on both the M-CHAT and	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **SES:** NR **Insurance status:** NR	**Performance characteristics (%)** Initial screen: M-CHAT: PPV: 36 (95% CI 31-40) M-CHAT-F: 74 (95% CI 68-80) Re-screening: M-CHAT: PPV: 38 (31 – 45) M-CHAT-F: 59 (50 – 68) **Other outcomes:** **Intermediate/Process outcomes:** ASD diagnoses

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Kleinman et al, 2008[3]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Spanish • Includes children from early intervention and pediatric referrals Note: See related publications Robins 2001,[5] Robins 2008,[2] Chlebowski 2013,[1] Pandey 2008[4]	At followup screening: 58.32 ± 8.66 **Sex, n (%):** Initial screening: M: 2003 (52.8) F: 1743 (46.0) NR: 47 (1.2) Re-screening: M: 800 (56.5) F: 602 (42.5) NR: 14 (1) **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Mental age:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **SES:** NR **Insurance status:** NR **Parental education:** NR **Marital status:** NR **Primary language spoken in home:** NR **Limited English proficiency:** NR	M-CHAT/F or physician/parent concerns **Total number available to screen:** Initial screen: 3793 Re-screening: 2469 **Total number screened:** Initial screen: 3793 Re-screening: 1416 **Number screening positive:** Initial screen: 385 Re-screening: M-CHAT: 201 M-CHAT/F: 124 **Number screening negative:** Initial screen: 3408 Re-screening: M-CHAT: 1215 M-CHAT/F: NR **Percent completing screening followup/diagnosis:** Initial screening: 5.4% (203/3793) Followup screening: n=131 (denominator unclear) **Diagnosis** **Diagnostic process:** ADOS, ADI-R, Mullen Scales of early Learning, the Vineland Adaptive Behavior Scales and CARS, DSM-IV Re-screening diagnostic pool included those who had received a diagnostic evaluation after initial screening due to screen failure or clinician/parent concern (n=120); failed the re-screener (n=8) ; or passed the rescreening but indicated they had been referred for possible ASD or other developmental delay in the time between screenings (n=3) **Who diagnosed:** Licensed clinical psychologist or developmental	**Parental education:** NR **Marital status:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Family context:** NR	missed at initial screening and discovered at re-screening: 7 **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Kleinman et al, 2008[3]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		pediatrician and a psychology doctoral student **Diagnoses, n:** ASD: 60 PDD-NOS:18 Asperger's: 2 Non-ASD diagnoses, n: Language delay: 13 Global developmental delay: 12 Other: 14 No diagnosis: 12 **Age at diagnosis, mean months ± SD:** Non-ASD: 55.87 ± 8.01 ASD: 52.17 ± 8.01		

Study Description Author/Year: Pandey et al, 2008[4]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:** • Excluded if received an ASD diagnosis before being screened with the M-CHAT • Severe sensory or motor disability (e.g., blindness or deafness) • Child's caregivers not fluent in English or Spanish • Includes children from early intervention and pediatric referrals Note: See related publications Robins 2001,[5] Kleinman 2008,[3]	**Population characteristics** **Population description:** Screening of toddlers at pediatric well-child visits or parent/clinician report of concerns **Race/ethnicity, n (%):** NR **Age at screening, mean months ± SD:** Age 16-23 months: 18.73 ± 1.23 Age 24-30 months: 25.12 ± 1.59 **Sex, n (%):** Age 16-23 months: M: 2368 (51) F: 2140 (47) NR: 74 (2) Age 24-30 months: M: 1216 (56) F: 949 (43) NR: 19 (1) **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Mental age:** NR	**Screening** **Instrument/method:** M-CHAT+ phone followup (M-CHAT/F) or parent/MD concern **Mode of administration:** Parent-rated, M-CHAT-F by phone **Setting:** Primary care **Administrator:** Research assistants **Scorer:** Research office **Study definition of positive screen:** Screening positive on 2 of 6 critical items or on 3 of 23 items overall on both the M-CHAT and M-CHAT/F or physician/parent concerns **Total number available to screen:** 6776 **Total number screened:** 6776 **Number screening positive:** 322 **Number screening negative:** 6454	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:**NR **Language development, mean ± SD:** NR **Family context:** NR	**Performance characteristics (%)** High risk=Parental/MD concerns, but no diagnosis Low risk= unselected well child visit sample Young: age 16-23 months Old: age 24-30 months **PPV by age group, %:** Total young: 66 Total old: 79 **PPV, % (95% CI):** Young/high risk: 79 (71-87) Old/high risk: 74 (65-83) Among high risk , young vs. old: p=NS Young/low risk: 28 (13-43) Old/low risk: 61 (44-78) Among low risk, younger vs. older subjects: p<0.01 **Other outcomes** **Intermediate/Process**

Study Description Author/Year: Pandey et al, 2008[4]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Chlebowski 2013,[1] Robins 2008[2]	**Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Percent completing screening followup/diagnosis:** 4% (270/6776) **Diagnosis** **Diagnostic process:** ADOS, ADI-R, Mullen Scales of early Learning, the Vineland Adaptive Behavior Scales and CARS, DSM-IV; included those who screened positive on M-CHAT/F and those flagged by pediatrician despite screen negative **Who diagnosed:** Licensed clinical psychologist or developmental pediatrician and a psychology doctoral student **Diagnoses, n:** ASD: 93 PDD-NOS: 91 Non-ASD: 66 No diagnosis: 16 Typical development:4 **Age at diagnosis, mean ± SD:** NR		**outcomes:** NR **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Robins, 2008[2]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** March 2005 – October 2007 **Inclusion/exclusion criteria:** • Excluded if received an ASD diagnosis before being screened with the M-CHAT • Severe sensory or motor disability (e.g., blindness or deafness)	**Population characteristics** **Population description:** Initial screen: Screening of toddlers at pediatric well-child visits or during early intervention intake or after referral from developmental pediatrician or psychologist Re-screening: children who passed initial screening or followup interview; also included participants from Robins et al, 2001 **Race/ethnicity, n (%):** N=1177 Caucasian: 779 (66.2)	**Screening** **Instrument/method:** M-CHAT+ phone followup (M-CHAT-F) at 2 time points **Mode of administration:** Parent-rated, M-CHAT-F by phone **Setting:** Primary care **Administrator:** Research assistants **Scorer:** Research office **Study definition of positive screen:** Screening positive on 2 of 6 critical items or on 3 of 23 items overall on both the M-CHAT and	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR	**Performance characteristics (%)** **PPV** M-CHAT: 5.8% M-CHAT-F: 57% **Other outcomes** **Intermediate/Process outcomes:** NR **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Robins, 2008[2]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
• Child's caregivers not fluent in English or Spanish • Includes children from early intervention and pediatric referrals Note: See related publications Robins 2001,[5] Kleinman 2008,[3] Chlebowski 2013,[1] Pandey 2008[4]	African American: 246 (20.9) Hispanic/Latino: 32 (2.7) Asian: 24 (2.1) Hawaiian/Pacific Islander: 2 (0.1) Other: 94 **Age at screening, mean months ± SD:** 20.92 ± 3,10 **At followup screening:** 58.32 ± 8.66 **Sex, n (%):** Initial screening: M: 2384 (49.7) F: 2280 (47.5) NR: 133 (2.8) Re-screening: M: 800 (56.5) F: 602 (42.5) NR: 14 (1) **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Mental age:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	M-CHAT-F or clinician suspicion (negative screens included those who declined phone interview or were excluded for other reasons) **Total number available to screen:** 4797 **Total number screened:** M-CHAT: 4797 M-CHAT-F: 362 **Number screening positive:** M-CHAT: 466 M-CHAT-F: 65 **Number screening negative:** M-CHAT: 4331 M-CHAT-F: 301 **Percent completing screening followup/diagnosis:** 1% (41/4797) **Diagnosis** **Diagnostic process:** ADOS, ADI-R, Mullen Scales of early Learning, the Vineland Adaptive Behavior Scales and CARS, DSM-IV **Who diagnosed:** Licensed clinical psychologist or developmental pediatrician, psychology doctoral student, an undergraduate research assistant to videotape the child **Diagnoses, n:** ASD: 21 Non-ASD: 20 **Age at diagnosis, mean months ± SD:** 23.56 ± 4.25	**Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Family context:** NR	**Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Robins et al, 2001[5]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:**	**Population characteristics** **Population description:** Initial screen: Screening of toddlers at pediatric well-child visits or during	**Screening** **Instrument/method:** M-CHAT+ phone followup (M-CHAT/F) at 2 time points	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR	**Performance characteristics (%)** **PPV:** of 58 children evaluated due to screen failure, 39 were

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Robins et al, 2001[5]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
• Excluded if received an ASD diagnosis before being screened with the M-CHAT • Severe sensory or motor disability (e.g., blindness or deafness) • child's caregivers not fluent in English or Spanish • Includes children from early intervention and pediatric referrals Note: See related publications Robins 2008,[2] Kleinman 2008,[3] Chlebowski 2013,[1] Pandey 2008[4]	early intervention intake **Race/ethnicity, n (%):** NR **Age at screening, mean months ± SD:** NR **Sex, n (%):** Initial screening: M: 2003 (52.8) F: 1743 (46.0) NR: 47 (1.2) Re-screening: M: 800 (56.5) F: 602 (42.5) NR: 14 (1) **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Mental age:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Mode of administration:** Parent-rated, M-CHAT-F by phone **Setting:** Primary care **Administrator:** Research assistants **Scorer:** Research office **Study definition of positive screen:** Screening positive on 2 of 8 critical items or on any 3 items overall on both the M-CHAT and M-CHAT/F **Total number available to screen:** 1293 **Total number screened:** 1293 **Number screening positive:** NR **Number screening negative:** M-CHAT: 1161 M-CHAT-F: 74 **Percent completing screening followup/diagnosis:** 4.5% (58/1293) **Diagnosis** **Diagnostic process:** ADOS, ADI-R, Mullen Scales of early Learning, the Vineland Adaptive Behavior Scales and CARS, DSM-IV **Who diagnosed:** Licensed clinical psychologist, graduate student clinician, and student who videotaped the session **Diagnoses, n:** ASD or PDD: 39 Non-ASD diagnoses, n: Language or global delays: 19 **Age at diagnosis, mean months:** ASD: 27.6 Non-ASD: 26.7 **Time to evaluation after screening, mean months:** 2.12	**Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:**NR **Language development, mean ± SD:** NR **Family context:** NR	diagnosed with ASD or PDD (67.2%) **Other outcomes** **Intermediate/Process outcomes:** ASD diagnoses missed at initial screening and discovered at re-screening: 7 **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Nygren et al, 2012[6]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Sweden **Study period:** Jan 2010 – December 2012 **Inclusion/exclusion criteria:** Children seen for 2.5-year-old checkup at child health center, any time during 2010	**Population characteristics** **Population description:** 3999 Swedish children screened during their 2.5-year well-child visit **Race/ethnicity, n (%):** **Age at screening, mean ± SD:** 2.5 years **Sex, n (%):** Female: 1912 (47.8) Male: 2087 (52.2) **Risk factors:** Family history of ASD: 12/48 (25%) **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Primary language spoken in home:** NR **Limited English proficiency:** NR	**Screening** **Instrument/method:** M-CHAT+ followup interview, JA-OBS **Mode of administration:** Interview **Setting:** Child health center **Administrator:** Nurse **Scorer:** Nurse **Study definition of positive screen:** Either a definitive failure on the M-CHAT or failure on ≥2 items on the JA-OBS or both **Total number available to screen:** 3999 **Total number screened:** 3999 **Number screening positive:** Total: 64 M-CHAT+interview: 36 JA-OBS: 40 Both: 48 **Number screening negative:** NR (estimated 3935 based on reported N screening positive) **Percent completing diagnostic followup, n:** 1.3 (54/3999) **Diagnosis** **Diagnostic process:** DSM-IV (based on all evidence obtained at the assessment and medical records) & ADOS Vineland Adaptive Behavior Scales; only screen positive children followed for diagnosis **Who diagnosed:** Clinicians **Diagnoses, n** **M-Chat +interview** ASD: 33/36 **JA-OBS** ASD: 37/40 **M-Chat+interview and/or JA-OBS** ASD: 43 /48 **Sub-types of ASD among the 48 children diagnosed, n:**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Among those diagnosed with ASD** **Race/ethnicity, n:** Parents of non-Swedish descent: 24/48 (50) One parent not of Swedish descent: 8/48 Both parents of Swedish descent: 16/48 **Age, mean ± SD:** NR **Sex, n (%):** Female: NR Male: NR **Risk factors:** Parental concern about child development prior to screening: 35/48 Family history of ASD: 12/48 **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development:** **Adaptive behavior, mean ± SD:** NR **Family context** SES: NR Insurance status: NR Parental education: NR Marital status: NR **Primary language spoken in**	**Performance characteristics, % (95% CI)** **M-CHAT+interview** (n=49): PPV: 91.7 (77.5-98.2) **JA-OBS** (n=48): PPV: 92.5 (79.6-98.4) **M-CHAT and/or JA-OBS** (n=51): PPV: 89.6 (77.3-96.5) **Other outcomes:** **Intermediate/Process outcomes:** NR **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Nygren et al, 2012[6]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Autistic disorder: 27 PDD-NOS: 21 **Non-ASD diagnoses, n** **M-Chat+interview+JA-OBS** Language Disorder:3 Typically Developing: 3 **Age at diagnosis, mean ± SD:** 2.5 ± 0.5 years	**home:** NR **Limited English proficiency:** NR	

Study Description Author/Year: Barbaro et al, 2011[7,8]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Australia **Study period:** September 2006 – 2008 **Inclusion/exclusion criteria:** • Children 8 to 24 months of age attending there well visits at a Maternal Child Health center in a 20-km radius of Melbourne University • "At risk" children referred by Maternal and Child Health nurse to the Social Attention and Communication Study (SACS) team • Parental consent given • Attendance at scheduled SACS assessment	**Population characteristics** **Population description:** From a population of 20,770 children monitored via the Social Attention and Communication Study (SACS) at maternal and child health centers, 216 children referred for further evaluation; consent obtained for 124 children; 110 children completed at least one SACS assessment. Included 10 12-month assessments; 46 18-month assessments; and 100 24-month assessments. **Race/ethnicity, n (%):** NR **Age at screening, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Screening** **Instrument/method:** Social Attention and Communication Study protocol **Mode of administration:** In person **Setting:** maternal-child health centers **Administrator:** Nurses **Scorer:** Nurses **Study definition of positive screen:** Items that were most relevant to ASD and developmentally appropriate for the age being monitored were considered "key" items. Children were considered at risk for an ASD only if they showed a "pattern" of failure on the items of interest, for example, by failing 3 of the 4 key items. **Total number available to screen:** 22168 **Total number screened:** 20770 **Number screening positive:** 216 **Number screening negative:** 20554 **Percent completing screening followup/diagnosis, (n):**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** Female: NR Male: NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean months ± SD:** 12 month assessment: Autistic Disorder: 13.7 ± 1.2 ASD: 12.7 ± 0.5 Developmental Delay/Language Delay: 15.0 18 month assessment: Autistic Disorder: 19.2 ± 1.0 ASD: 19.1 ± 1.2 Developmental Delay/Language Delay: 19.9 ± 1.6 24 month assessment Autistic Disorder: 25.2 ± 1.6 ASD: 25.6 ± 2.2	**Performance characteristics** **SACS assessment process:** **PPV (%):** at 12 months: 90 at 18 months: 79 at 24 months: 81 **Other outcomes** **Intermediate/Process outcomes:** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Barbaro et al, 2011[7,8]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Screen positive: 0.5 (110) **Diagnosis** **Diagnostic process:** At 12 and 18 mo: investigator administered: Mullen Scales of Early Learning, Early Social and Communication Scales; imitation/ name call/spontaneous play tasks (+empathy at 18 mo only); CHAT-23 (at 18 mo only) Parent questionnaires: Infant Toddler Checklist-CSBS-DP; Early Development Interview; CHAT-23 (18 mo only) At 24 mo: investigator administered: Mullen Scales of Early Learning; ADOS; Imitation/ empathy tasks Parent questionnaires: ADI-R **Who diagnosed:** SACS investigators **Diagnoses, n** **ASD:** 50 **Autistic Disorder:** 39 **Developmental delay or other language delay:** 20 **Typically developing:** 1 **Age at diagnosis, mean ± SD** NR	Developmental Delay/Language Delay: 25.8 ± 2.7 **Sex, n (%):** 12 month assessment: Female Autistic Disorder: 1 ASD: 1 Developmental Delay/Language Delay: 0 Male Autistic Disorder: 2 ASD: 5 Developmental Delay/Language Delay: 1 18 month assessment Female Autistic Disorder: 4 ASD: 1 Developmental Delay/Language Delay: 3 Male Autistic Disorder: 12 ASD: 20 Developmental Delay/Language Delay: 5 24 month assessment Female Autistic Disorder: 10 ASD: 8 Developmental Delay/Language Delay: 6 Male Autistic Disorder: 27 ASD: 34 Developmental Delay/Language Delay: 8 **Risk factors:** NR **IQ, mean ± SD:** NR Overall mental age, months 12 months Autistic Disorder: 10.6 ± 2.6 ASD:10.3 ± 2.2	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Barbaro et al, 2011[7,8]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			DD/LD: 12.3 18 months Autistic Disorder: 13.2 ± 1.9 ASD: 14.8 ± 1.8 Developmental Delay/Language Delay: 15.4 ± 1.5 24 months Autistic Disorder: 15.1 ± 2.5 ASD: 18.5 ± 2.9 Developmental Delay/Language Delay: 19.5 ± 3.3 Nonverbal age, months 12 months: Autistic Disorder: 12.0 ± 4.3 ASD: 11.5 ± 2.2 Developmental Delay/Language Delay: 16.5 18 months Autistic Disorder: 17.3 ± 2.3 ASD: 17.5 ± 2.9 Developmental Delay/Language Delay: 17.4 ± 1.9 24 months Autistic Disorder: 19.1 ± 2.9 ASD: 21.4 ± 2.7 Developmental Delay/Language Delay: 21.3 ± 3.6 **T score, visual reception** 12 months Autistic Disorder: 29.7 ± 6.5 ASD: 37.0 ± 9.2 Developmental Delay/Language Delay: 47.0 18 months Autistic Disorder: 37.6 ± 8.5 ASD: 39.0 ± 9.1 Developmental Delay/Language Delay:37.0 ± 7.5 24 months Autistic Disorder: 30.9 ± 7.6 ASD: 35.8 ± 8.1 Developmental Delay/Language	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Barbaro et al, 2011[7,8]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			Delay: 36.6 ± 9.9 **T score, fine motor:** 12 months Autistic Disorder: 43.7 ± 20.1 ASD: 44.7 ± 10.8 Developmental Delay/Language Delay: 53.0 18 months Autistic Disorder: 44.0 ± 10.9 ASD: 44.9 ± 11.8 Developmental Delay/Language Delay: 42.1 ± 6.5 24 months Autistic Disorder: 36.0 ± 11.1 ASD: 40.7 ± 9.0 Developmental Delay/Language Delay: 37.8 ± 11.0 **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** Verbal age 12 months Autistic Disorder: 9.2 ± 2.9 ASD: 9.0 ± 2.5 Developmental Delay/Language Delay: 8.0 18 months Autistic Disorder: 9.2 ± 2.2 ASD: 12.0 ± 1.5 Developmental Delay/Language Delay: 13.4 ± 2.0 24 months Autistic Disorder: 11.0 ± 2.7 ASD: 15.8 ± 4.1 Developmental Delay/Language Delay: 17.6 ± 3.5 **T score, receptive language** 12 months Autistic Disorder: 28.5 ± 5.2 ASD: 30.8 ± 7.3 Developmental Delay/Language Delay: 20.0	

Study Description Author/Year: Barbaro et al, 2011[7,8]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			18 months Autistic Disorder: 20.6 ± 2.5 ASD: 24.4 ± 5.0 Developmental Delay/Language Delay: 25.5 ± 4.5 24 months Autistic Disorder: 20.3 ± 1.6 ASD: 26.3 ± 9.2 Developmental Delay/Language Delay: 32.2 ± 10.4 **T score, expressive language** 12 months Autistic Disorder: 31.3 ± 3.8 ASD: 38.2 ± 10.6 Developmental Delay/Language Delay: 28.0 18 months Autistic Disorder: 26.3 ± 4.6 ASD: 31.4 ± 4.9 Developmental Delay/Language Delay: 35.0 ± 4.3 24 months Autistic Disorder: 23.9 ± 4.1 ASD: 31.7 ± 7.4 Developmental Delay/Language Delay: 32.5 ± 6.5 **Adaptive behavior, mean ± SD:** NR **Family context:** NR	

Study Description Author/Year: Canal-Bedia et al, 2011a[9]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed
Country: Spain **Study period:** October 2005 – April 2008 **Inclusion criteria:** • Age 18-36 months • Resident in geographic area of interest during study period	**Population characteristics** **Population description:** Unselected children at well child visit **Race/ethnicity, n (%):** NR **Age at screening, range:** Unselected population: 18-24 months **Sex, n (%):** Female: 949 (46)	**Screening** **Instrument/method:** M-CHAT with followup interview (Spanish translation) **Mode of administration:** In-person during well child visit, telephone followup for those screening positive **Setting:** Clinic **Administrator:**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Canal-Bedia et al, 2011a[9]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed
	Male: 1106 (54) **Risk factors:** None **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	Initial screening: Primary care pediatricians and nurses Phone followup: Psychologist with child development training **Scorer:** Initial screening: Central research team Phone followup: Psychologist conducting the phone interview **Study definition of positive screen:** Failure of 3/23 items or 2/6 critical items **Total number available to screen:** NR **Total number screened:** 2055 **Number screening positive:** Initial screening: 336 Phone followup: 31 **Number screening negative:** Initial screening: 1719 Phone followup: 305 **Percent completing screening phone followup:** Screen positive: 100 Screen negative: 100 **Percent completing diagnosis:** Screen Positive: 100 **Diagnosis** **Diagnostic process:** Team of clinicians using diagnostic algorithm based on recommendations of American Academy of Neurology and Child Neurology Society plus VABS, Merrill-Palmer revised scales, ADOS-G module 1 with cases classified according to DSM-IV-TR **Who diagnosed:** Psychologists and neurologists **Diagnoses, n:** ASD: 6 Non-ASD: 25 **Age at diagnosis, range:** 18-24 months	**Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Canal-Bedia et al, 2011b[9]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Spain **Study period:** October 2005 – April 2008 **Inclusion criteria:** • Age 18-36 months or for children at high risk, maximum chronological age of 48 months with developmental age of 18-24 months • Resident in geographic area of interest during study period	**Population characteristics** **Population description:** Unselected children at well child visit and children considered at high risk for ASD recruited from early intervention or child psychiatry units **Race/ethnicity, n (%):** NR **Age at screening, range:** Unselected population: 18-24 months High-risk population: NR **Sex, n (%):** Unselected population Female: 1163 (47) Male: 1254 (51) High-risk population Female: 14 (.56) Male: 49 (2) **Risk factors:** Attending early intervention or child psychiatry units, n=63 **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Primary language spoken in home:** NR **Limited English proficiency:** NR	**Screening** **Instrument/method:** M-CHAT (with followup interview) (Spanish translation) **Mode of administration:** In-person during well child visit, phone followup for those screening positive **Setting:** Clinic **Administrator:** Initial screening: Primary care pediatricians and nurses Phone followup: Psychologist with child development training **Scorer:** Initial screening: Central research team Phone followup: Psychologist conducting the phone interview **Study definition of positive screen:** Failure of 3/23 items or 2/6 critical items **Total number available to screen:** NR **Total number screened:** 2480 **Number screening positive:** Initial screening: 429 Phone follow-up: 86 **Number screening negative:** Initial screening: 2051 Phone interview: 343 **Percent completing screening telephone followup:** Screen Positive: 100 Screen Negative: 100 **Percent completing diagnostic followup:** Screen Positive: 100 **Diagnosis** **Diagnostic process:** Team of clinicians using diagnostic algorithm based on recommendations of American Academy of Neurology and	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context** SES: NR Insurance status: NR Parental education : NR Marital status: NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Performance characteristics** **Initial screen+phone followup, % (95% CI)** **PPV:** 35 (23-46) **AUC:** 99 **Other outcomes** **Intermediate/Process outcomes:** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Canal-Bedia et al, 2011b[9]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Child Neurology Society plus VABS, Merrill-Palmer revised scales, ADOS-G module 1 with cases classified according to DSM-IV-TR **Who diagnosed:** Psychologists and neurologists **Diagnoses, n ASD:** Unselected population: 4 High-risk population: 19 **Non-ASD diagnoses, n:** Typical development:4 Non-ASD developmental disorder: 59 **Age at diagnosis, range:** Unselected population: 18-24 months High-risk population: NR		

Study Description Author/Year: Inada et al, 2011[10]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Japan **Study period:** Reliability: NR Concurrent Validity: December 2008 – December 2009 Discriminant Validity: April 2005 – March 2007 **Inclusion/exclusion criteria:** • Validity samples: Discriminant: 18 month olds attending health checkup in Munakata city, Japan between April 2005 and March 2007	**Population characteristics** **Population description:** *Discriminant Validity Sample* Collected from children at free health checkup visits for 18 month olds **Race/ethnicity, n (%):** NR **Age at screening, mean ± SD:** *Discriminant Validity Sample: NR* (but collected at 18 month well-baby visit) **Sex, n (%):** *Discriminant Validity Sample* Female: 575 (48.4) Male: 612 (51.6) **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Primary language spoken in home:** NR **Limited English proficiency:** NA	**Screening** **Instrument/method:** Japanese M-CHAT 23-item version, M-CHAT 9-item version (short) **Mode of administration:** Parent-rated **Setting:** Clinic **Administrator:** NR **Scorer:** Research team **Study definition of positive screen:** 23 item M-CHAT: failing 2/23 items 9 item M-CHAT: failing 1/9 items **Total number available to screen:** Discriminant Validity: 1187 **Total number screened:** Discriminant Validity: 1187 **Number screening positive** **Cut-off value ≥2:** 23 items:148	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** 24 months ± 0.7 months **Sex, n (%):** Female ASD: 4 Non-ASD: 571 Male ASD: 16 Non-ASD: 596	**Performance characteristics** ***23 item M-CHAT*** ***Cut off Score ≥1*** **Sensitivity:** 80 **Specificity:** 60.9 **PPV:** 3.4 **NPV:** 99.4 ***Cut off Score ≥2*** **Sensitivity:** 75 **Specificity:** 89.3 **PPV:** 10.7 **NPV:** 99.5 ***Cut off Score ≥3*** **Sensitivity:** 55 **Specificity:** 96.1 **PPV:** 19.3 **NPV:** 99.2 ***Cut off Score ≥4*** **Sensitivity:** 35 **Specificity:** 97.8 **PPV:** 21.2

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Inada et al, 2011[10]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		9 items: 60 **Number screening negative** **Cut-off value ≥2:** 23 items: 1039 9 items: 1127 **Percent completing screening followup/diagnosis:** **Screen Positive+Screen Negative:** 100 **Diagnosis** **Diagnostic process:** ***Discriminant Validity Sample*** Children were diagnosed based on developmental history, clinical assessment and based on DSM-IV-TR criteria **Who diagnosed:** Research team, including 2 psychiatrists and psychologist **Diagnoses, n** ***Discriminant Validity Sample*** **Among those screening positive (cut-off value ≥2):** Using 23 items: ASD: 15 Non-ASD: 133 Using 9 items: ASD:11 Non-ASD:49 **Among those screening negative (cut-off value ≥2):** Using 23 items: ASD: 5 Non-ASD: 1034 Using 9 items: ASD: 9 Non-ASD:1118 **Age at diagnosis, mean ± SD:** 24 months ± 0.7 months	**Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**NPV:** 98.9 ***9 item M-CHAT*** ***Cut off Score ≥1*** **Sensitivity:** 65 **Specificity:** 88.5 **PPV:** 8.8 **NPV:** 99.3 ***Cut off Score ≥2*** **Sensitivity:** 55 **Specificity:** 95.8 **PPV:** 18.3 **NPV:** 99.2 ***Cut off Score ≥3*** **Sensitivity:** 40 **Specificity:** 98.1 **PPV:** 26.7 **NPV:** 99.0 ***Cut off Score ≥4*** **Sensitivity:** 20 **Specificity:** 99.0 **PPV:** 25.0 **NPV:** 98.6 **Other outcomes** **Intermediate/Process outcomes:** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Miller et al, 2011[11]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:** Toddlers aged 14-30 months presenting to the community based pediatric practice during the 6 month screening period	**Population characteristics** **Population description:** Children attending any visit during the study period (well-child, sick, followup or injection visit) **Race/ethnicity, n (%):** N=524 Black: 17 (3.2) Pacific Islander: 21 (4.0) Asian: 22 (4.2) Native American: 11 (2.0) White, Non-Hispanic: 287 (55) Other: 1 (0.2) Hispanic: 167 (32) **Age at screening, mean ± SD:** Range: 14-24 months **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Primary language spoken in home, n (%):** English: 721 (91) Spanish: 75 (9) **Limited English proficiency:** NR	**Screening** **Instrument/method:** M-CHAT + followup interview or Infant Toddler Checklist (ITC) + followup interview **Mode of administration:** Parent-rated **Setting:** Primary Care **Administrator:** self-administered by caregiver; followup phone call by research staff **Scorer:** Research Staff **Study definition of positive screen:** ITC: Score below 10th percentile on Social, Symbolic or Total Score M-CHAT: Failure of 2+ critical items or 3+ of any item **Total number available to screen:** 990 **Total number screened:** 796 **Number screening positive:** At initial questionnaire: 192 At followup phone call: 47 **Number screening negative:** At initial questionnaire: 544 At followup phone call: 98 **Percent completing screening followup/diagnosis:** Overall: 30 /796 (4%) Screening Positive: 28/192 (14.6%) Screening Negative: 2 / 544 (0.4 %) **Diagnosis** **Diagnostic process:** Parents of children who screened positive on either screener, or were referred by a clinician due to concerns (regardless of screen outcome) were called. If positive M-CHAT, M-CHAT followup interview was completed by a graduate student. If positive ITC, unspecified relevant items were repeated and discussed	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** **ASD** Black: 1 (8) Hispanic: 4 (30.7) White, Non-Hispanic: 8 (61.5) **Non-ASD** Black: 0 (0) Hispanic: 4 (23.5) White, Non-Hispanic: 11 (64.7) Pacific Islander: 2 (11.8) **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Performance characteristics (%)** M-CHAT/ITC: Sensitivity: NR PPV: 32 (8/25) **Other outcomes:** **Intermediate/Process outcomes:** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Miller et al, 2011[11]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		over the phone. If clinician referral, M-CHAT and ITC were repeated over the phone. Children who had a confirmed positive screen over the phone were invited for diagnostic evaluation, including ADOS-T, Mullen, and Vineland. **Who diagnosed:** Research reliable clinician **Diagnoses, n** Among screen positives (n=25): ASD: 8 Non-ASD diagnoses, n: Developmental delay: 16 Typically Developing: 1 **Among screen negatives (n=2):** ASD: 2 Non-ASD: 0 **Age at diagnosis:** NR		

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** NR **Inclusion/exclusion criteria:** Infants attending 1-year checkup at various pediatrician offices in San Diego county, California	**Population characteristics** **Population description:** Children at well-child check visits; 184 children who screened positive at 12-month well baby visits; 41 control children **Race/ethnicity, n (%):** NR **Age at screening, mean ± SD:** Mean = 12.54 months Range = 10.08 to 15.97 months **Sex, n (%):** NR **Risk factors:** NR. **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Primary language spoken in home:** NR	**Screening** **Instrument/method:** CSBS-DP-IT **Mode of administration:** Parent-rated **Setting:** Primary care **Administrator:** checklist is given to parents by receptionist at primary care office; parent self-administers **Scorer:** Medical Staff **Total number available to screen:** 10479 **Total number screened:** 10479 **Number completing screening followup/diagnosis:** Screen Positive: 184 Screen Negative: 41 **Number screening positive:** 1318 (12.5%)	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR: **Risk factors:** NR **IQ, mean ± SD:** Screen Positive: NR Screen Negative: 111.2 ± 11.7 **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **At Earliest Age** **Race/ethnicity, n (%):** NR **Sex, n (%):** NR **Risk factors:** NR	**Performance characteristics (%)** Sensitivity: 100 Specificity: 47 PPV: 75 NPV: 100 **Other outcomes** **Intermediate/Process outcomes:** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** Children began treatment on average at 17 months **Mean amount of intervention (hours/week):** ASD: 11.5

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
	Limited English proficiency:	**Number screening negative:** 9179 (87.5%) **Diagnosis** **Diagnostic process:** Developmental evaluation including ADOS-T and Mullen. Children were followed every 6 months until 3 years of age. **Who diagnosed:** Psychologist **Diagnoses, n** ASD: 32 Initial ASD diagnosis, removed at follow-up: 5 Language Delay: 56 Developmental Delay: 9 Other Diagnosis: 36 False Positive: 46 Typically Developing: 41 **Age at diagnosis (latest), mean ± SD months:** ASD: 34.7 ± 8.8 Initial ASD, removed at follow-up: 32.0 ± 8.9 Language Delay: 25.9 ± 9.2 Developmental Delay: 33.4 ± 9.8 Other: 25.7 ± 8.6 False-Positive Screen: 25.4 ± 8.7 Typically Developing: 25.5 ± 10.3 **Number of Visits, mean ± SD** ASD: 6.0 ± 2.5 Initial ASD, removed at followup: 4.8 ± 2.2 Language Delay: 4.4 ± 2.4 Developmental Delay: 5.8 ± 2.9 Other: 4.3 ± 2.1 False-Positive Screen: 4.6 ± 2.3 Typically Developing: 4.3 ± 2.7 **% of sample tracked until 24 months old or older** ASD: 97% Initial ASD, removed at followup: 80%	**Age at IQ Testing, mean ± SD:** ASD: 17.8 ± 3.9 Language Delay: 15.3 ± 3.1 Developmental Delay: 17.0 ± 3.9 Other: 14.5 ± 2.5 Initial ASD, removed at followup: 15.9 ± 4.4 False-Positive Screen: 13.7 ± 1.4 Typically Developing: 14.5 ± 2.4 **IQ (Early Learning Composite), mean ± SD:** MSEL; Standard Score ASD: 75.5 ± 15.4 Language Delay: 90.4 ± 11.6 Developmental Delay: 63.3 ± 11.1 Other: 96.3 ± 12.0 Initial ASD, removed at followup: 92.6 ± 9.1 False-Positive Screen: 106.1 ± 11.9 Typically Developing: 111.2 ± 11.7 **IQ (Visual Reception), mean ± SD:** MSEL; T-Scores ASD: 43.0 ± 10.2 Language Delay: 51.0 ± 10.2 Developmental Delay: 30.5 ± 7.3 Other: 52.8 ± 9.5 Initial ASD, removed at follow-up: 51.6 ± 7.2 False-Positive Screen: 55.0 ± 8.5 Typically Developing: 56.5 ± 8.2 **Language development (Receptive Language), mean ± SD:** MSEL; T-Score	Language Delay: 1.9 Developmental Delay: 8.5 **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Language Delay: 54% Developmental Delay: 78% Other: 61% False-Positive Screen: 61% Typically Developing: 56% **% of sample tracked until 32 months old or older** ASD: 72% Initial ASD, removed at followup: 60% Language Delay: 30% Developmental Delay: 67% Other: 26% False-Positive Screen: 28% Typically Developing: 42%	ASD: 29.1 ± 11.4 Language Delay: 39.6 ± 8.9 Developmental Delay: 23.4 ± 3.1 Other: 40.7 ± 8.4 **Initial ASD, removed at followup:** 34.8 ± 5.5 False-Positive Screen: 46.7 ± 8.3 Typically Developing: 53.4 ± 11.0 **Language development (Expressive Language), mean ± SD:** MSEL; T-Score ASD: 32.2 ± 11.7 Language Delay: 37.0 ± 7.3 Developmental Delay: 26.7 ± 8.0 Other: 45.3 ± 8.0 Initial ASD, removed at follow-up: 41.6 ± 8.1 False-Positive Screen: 51.6 ± 9.6 Typically Developing: 53.3 ± 7.4 **Age at Earliest ADOS, mean ± SD:** ASD: 17.9 ± 3.8 Language Delay: 15.6 ± 3.1 Developmental Delay: 16.2 ± 3.1 Other: 14.7 ± 2.5 Initial ASD, removed at follow-up: 16.2 ± 4.8 False-Positive Screen: 14.0 ± 1.5 Typically Developing: 15.0 ± 2.3 **ASD severity, mean ± SD:** ADOS Total Score ASD: 16.8 ± 6.0 Language Delay: 6.5 ± 4.3	

Study Description Author/Year: Pierce et al, 2011[12]	**Baseline/Prescreening Characteristics**	**Screening and Diagnosis**	**Characteristics of Population Screened and/or Diagnosed**	**Screening Outcomes**
			Developmental Delay: 12 ± 7.2 Other: 6.2 ± 4.1 Initial ASD, removed at follow-up: 11.6 ± 5.0 False-Positive Screen: 3.3 ± 3.0 Typically Developing: 2.5 ± 2.1 **ASD severity, mean ± SD:** ADOS Social Affect Score ASD: 13.8 ± 4.8 Language Delay: 5.3 ± 3.3 Developmental Delay: 9.6 ± 6.4 Other: 4.3 ± 3.2 Initial ASD, removed at follow-up: 9.8 ± 4.1 False-Positive Screen: 3.0 ± 2.9 Typical: 2.1 ± 1.8 **ASD severity, mean ± SD:** ADOS RR Score ASD: 2.9 ± 1.8 Language Delay: 1.2 ± 1.4 Developmental Delay: 2.3 ± 2 Other: 1.8 ± 2.1 Initial ASD, removed at follow-up: 1.8 ± 1.6 False-Positive Screen: 0.3 ± 0.6 Typically Developing: 0.5 ± 1.0 **Adaptive behavior, mean ± SD:** NR **Family context:** NR **At Latest Age** **Race/ethnicity, n (%):** NR **Sex, n (%):** NR **Risk factors:** NR **Age at Latest IQ Testing, mean months ± SD:** ASD: 34.1 ± 8.9 Language Delay: 25.1 ± 9.0 Developmental Delay: 32.4 ± 11.2	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			Other: 24.9 ± 8.8 Initial ASD, removed at follow-up: 31.7 ± 9.3 False-Positive Screen: 24.6 ± 8.9 Typically Developing: 23.8 ± 10.2 **IQ (Early Learning Composite), mean ± SD:** MSEL; Standard Score ASD: 78.6 ± 17.5 Language Delay: 98.8 ± 17.4 Developmental Delay: 58.8 ± 22.1 Other: 98.6 ± 18.1 Initial ASD, removed at follow-up: 99.8 ± 8.1 False-Positive Screen: 110.9 ± 13.3 Typically Developing: 111.8 ± 13.6 **IQ (Visual Reception), mean ± SD:** MSEL; T-Scores ASD: 41.0 ± 11.6 Language Delay: 54.7 ± 12.6 Developmental Delay: 31.8 ± 12.2 Other: 51.8 ± 12.3 Initial ASD, removed at follow-up: 50.6 ± 3.9 False-Positive Screen: 57.3 ± 8.7 Typically Developing: 56.9 ± 11.1 **Language development (Receptive Language), mean ± SD:** MSEL; T-Score ASD: 35.5 ± 12.1 Language Delay: 45.7 ± 11.3 Developmental Delay: 30.4 ±	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			9.4 Other: 46.1 ± 11.1 Initial ASD, removed at follow-up: 46.8 ± 10.7 False-Positive Screen: 51.8 ± 9.2 Typically Developing: 54.3 ± 11.7 **Language development (Expressive Language), mean ± SD:** MSEL; T-Score ASD: 37.6 ± 14.6 Language Delay: 43.2 ± 11.2 Developmental Delay: 30.9 ± 8.3 Other: 47.6 ± 10.0 Initial ASD, removed at follow-up: 54.0 ± 6.2 False-Positive Screen: 51.8 ± 10.4 Typically Developing: 53.1 ± 8.3 **Age at Latest ADOS, mean ± SD:** ASD: 34.2 ± 8.8 Language Delay: 25.1 ± 8.8 Developmental Delay: 32.6 ± 11.2 Other: 24.7 ± 9.1 Initial ASD, removed at follow-up: 32.1 ± 10.3 False-Positive Screen: 24.9 ± 8.7 Typically Developing: 24.6 ± 9.9 **ASD severity, mean ± SD:** ADOS Total Score ASD: 16.9 ± 5.2 Language Delay: 4.1 ± 3.5 Developmental Delay: 6.9 ± 4.8 Other: 4.8 ± 3.0 Initial ASD, removed at follow-up: 5.3 ± 1.0	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Pierce et al, 2011[12]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
			False-Positive Screen: 2.6 ± 2.5 Typically Developing: 1.9 ± 2.1 **ASD severity, mean ± SD:** ADOS Social Affect Score ASD: 13.4 ± 4.6 Language Delay: 3.6 ± 2.9 Developmental Delay: 4.6 ± 3.4 Other: 4.0 ± 2.4 Initial ASD, removed at follow-up: 4.8 ± 1.9 False-Positive Screen: 2.3 ± 2.3 Typically Developing: 1.7 ± 1.8 **ASD severity, mean ± SD:** ADOS RR Score ASD: 3.5 ± 1.6 Language Delay: 0.6 ± 1.0 Developmental Delay: 2.3 ± 1.8 Other: 0.8 ± 1.6 Initial ASD, removed at follow-up: 0.5 ± 1.0 False-Positive Screen: 0.3 ± 0.7 Typically Developing: 0.2 ± 0.5 **Adaptive behavior, mean ± SD:** NR **Family context:** NR	

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Honda et al, 2009[13]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Japan **Study period:** 1989 to 1992 **Inclusion/exclusion criteria:** Children undergoing 18-month checkup in Yokohama	**Population characteristics** **Population description:** Children undergoing 18-month checkup **Race/ethnicity, n (%):** NR **Age at screening, mean ± SD:** NR (18 month well child visit) **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Screening** **Instrument/method:** YACHT 18 (Young Autism and other developmental disorders CHeckup Tool) used at 18 month routine health checkup **Mode of administration:** In person **Setting:** Pediatric well-baby visits **Administrator:** Nursing staff **Scorer:** Nursing staff **Study definition of positive screen:** NR; Public health nurse assessment of all screening data **Total number available to screen:** 3036 **Total number screened:** 2,814 **Number screening positive:** 402 **Number screening negative:** 2412 **Percent completing screening followup/diagnosis:** 0.82% (23/2814) **Diagnosis** **Diagnostic process:** NR **Who diagnosed:** NR **Diagnoses, n** **Among all participants:** ASD: 5 PDD-NOS: 10 Learning disorders: 1 AD/HD: 6 Mental retardation: 4 Cerebral palsy: 1 Erb's palsy: 1 **Among those screening positive:** Total: 19 ASD: 3 PDD-NOS: 8 Learning disorders: 1 AD/HD: 5	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Performance characteristics (%)** YACHT (18 months, autistic disorder and PDD) Sensitivity: 79 Specificity: 11 PPV: 58 NPV: 25 For any developmental disorder: Specificity: 86.3 (extraction stage: 2408/2791) 100 (refinement stage) Among IQ ≥70: Sensitivity: 86.7 (13/15) Among IQ ≤69: Sensitivity: 75 (6/8) **Other outcomes:** **Intermediate/Process outcomes:** **Age at diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Honda et al, 2009[13]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		Mental retardation: 2 Cerebral palsy: 0 Erb's palsy: 0 **Among those screening negative:** Total: 4 ASD: 2 PDD-NOS: 1 Learning disorders: 0 AD/HD: 1 Mental retardation: 0 Cerebral palsy: 0 Erb's palsy: 0 **Age at diagnosis, mean ± SD:** NR		

Study Description Author/Year: Wetherby et al, 2008[14]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: US **Study period:** June 1997 to November 2003 **Inclusion/exclusion criteria:** • Between 4 and 10 years old • Born between June 1997 and November 2003	**Population characteristics** **Population description:** General population recruited from childcare and health care agencies **Race/ethnicity, n (%)** Caucasian, 3193 (59.3) African American, 1626 (30.2) Hispanic, 167 (3.1) Asian, 124 (2.3) Other, 275 (5.1) **Age at screening, mean ± SD:** 13.5 months ± 4.7 (at first screen) **Sex, n (%):** Female: 2633(48.9) Male: 2752 (51.1) **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context** **Parental education** Years of Education Completed	**Screening** **Instrument/method:** Infant Toddler Checklist **Mode of administration:** Parent-rated, independently **Setting:** Home, clinic **Administrator:** NR **Scorer:** NR **Study definition of positive screen:** Score in the bottom 10th percentile on Social Composite, Symbolic composite or Total score on ITC *OR* bottom 10th percentile on speech composite on 2 consecutive ITCs. **Total number available to screen:** 5385 **Total number screened:** 5385 **Number screening positive:** 969 (18%) **Number screening negative:** 4416 (82%) **Percent completing screening**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Performance characteristics** **Sensitivity:** 93.3 **Other outcomes** **Intermediate/Process outcomes** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Wetherby et al, 2008[14]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
	Mothers: 14.3 ± 2.4 Fathers: 14.3 ± 2.7 **Parental age at birth of child** Mothers: 28.2 ± 6.2 Fathers: 30.9 ± 7.0 **Primary language spoken in home:** NR **Limited English proficiency:** NR	**followup/diagnosis:** Among those screening positive: NR Among those screening negative: NR **Diagnosis** **Diagnostic process:** Behavior Profile (CSBS), Mullen Vineland, ADOS, Social Communication Questionnaire **Who diagnosed:** NR **Diagnoses, n** **ASD:** 60 (56 of these had screened positive) **Non-ASD diagnoses:** NR **Age at diagnosis, range:** 4 to 10 years		

Study Description Author/Year: Dietz et al, 2007[15]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Netherlands **Study period:** October 1999 to April 2002 **Inclusion criteria:** **C**hildren screened at well-baby visit at age 14-15 months **Exclusion criteria:** • Parents did not cooperate (n=52) • Screening lists incomplete or physician unable to prescreen child, usually due to language barrier (n=399) Note: Population may overlap with Dietz et al[16]	**Population characteristics** **Population description:** Children screened at well-baby visit at age 14-15 months **Race/ethnicity, n (%):** NR **Age at screening:** Screen 1:14-15 months Screen 2: 16 ± 2months (from population screening) 27 ± 6 months (from surveillance) **Sex, %: (n=173)** Female: 25 Male: 75 **Risk factors:** NR **DQ, n:** By compliance status: Early: 69.94 ± 17.96 Late: 72.54 ± 18.88 Noncompliance: 89.19 ± 14.95	**Screening** **Instrument/method:** 2 stage screening: 4- & 14-item ESAT + developmental surveillance. Children initially prescreened with 4-item ESAT by their doctor; parents of those who screened positive or those identified by surveillance were invited to have a 1.5-hr home visit with a child psychologist who conducted the 14-item ESAT. **Mode of administration:** In-person **Setting:** Clinic; home **Administrator:** Trained child psychologist **Scorer:** Administrator **Study definition of positive**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):**NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean**	**Performance characteristics (%):** 4-item ESAT + developmental surveillance: PPV: 29 14-item ESAT: PPV: 41.1 **Other outcomes** **Intermediate/Process outcomes** **Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Study Description Author/Year: Dietz et al, 2007[15]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
	Language development, mean ± SD: NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**screen:** Prescreen: negative answer by parent on 1/4 ESAT items; screening: ≥3 negative answers in 14-item ESAT by both parent and child psychologist **Total number available to screen:** 31,724 Screen 1 (4-item ESAT): Total screened: 31,724 Number screening positive: 370 Number screening negative: 31,354 Screen 2 (14-item ESAT) Total screened: 364 (255 of screen 1 positives & 109 through surveillance) Number screening positive: 173 Number screening negative: 191 **Percent completing screening followup/diagnosis:** 17.8% (66/370) of screen 1 positives 81.5 % (141/173) of screen 2 positives **Diagnosis** **Diagnostic process:** 5 tests over 5 weeks including standardized parent interview, developmental history, cognitive testing (Bayley Scales, Mullen Scales of Early Learning, Psychoeducational Profile-Revised), Vineland Social-Emotional Early Childhood Scales, standardized behavior observation using ADOS or ADOS-G, pediatric examination and medical workup. Subset of patients undergoing screening with SCQ had diagnostic workup including ADI-R. **Who diagnosed:** Experienced child psychiatrist made a clinical	**± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	

Study Description Author/Year: Dietz et al, 2007[15]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		judgment about whether the child would meet DSM-IV criteria of different diagnoses by age 3.5 years. **Diagnoses, n** Screen1: Total diagnosed: 66/255 complied (ASD=19, non-ASD= 47) Screen 2: Total diagnosed: 141/173 (ASD=58; non-ASD=83) **Age at diagnosis, mean ± SD:** 3.5 years		

Study Description Author/Year: VanDenHeuvel et al, 2007[17]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: UK **Study period:** NR **Inclusion/exclusion criteria:** • 18-month-olds • Attending well check visit	**Population characteristics** **Population description:** 2117 infants attending the routine 18-month developmental assessment **Race/ethnicity, n (%):** NR **Age at screening, mean ± SD:** 18 months **Sex, n (%):** Female: 1029 (49%) Male: 1088 (51%) **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context** **SES:** Representative of the area social class distribution of based on 2002 census data **Primary language spoken in home:** NR **Limited English proficiency:** NR	**Screening** **Instrument/method:** Checklist for Autism in Toddlers (CHAT), 14-item interviewer-administered instrument divided into 2 sections: Section A includes 9 items administered to the parent and Section B includes 5 items based on interviewer observations of the infant **Mode of administration:** Interview /observation **Setting:** Clinic **Administrator:** Public health nurses **Scorer:** Public health nurses **Study definition of positive screen:** Each completed CHAT was scored into 1 of 3 categories: high, medium, or low risk for autism based on a standard scoring system. If an infant scored medium or high risk for autism at the first administration, a second screening	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Performance characteristics (%):** PPV: 58.3 (7/12) **Other outcomes:** **Intermediate/Process outcomes** **Time between screening and diagnosis:** NR **Time between screening and intervention:** NR **Health-related outcomes:** NR

Study Description Author/Year: VanDenHeuvel et al, 2007[17]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		was administered approximately 1 month later. **Total number available to screen:** 2684 **Total number screened:** 2117 **Number screening positive:** 29 **Number screening negative:** 2088 **Percent completing screening followup:** 24.1 (7/29) **Percent completing diagnosis:** Overall: 0.57 (12/2117) Screen Positive: 41.4 (12/29) Screen Negative: 0 **Diagnosis** **Diagnostic process:** Diagnostic assessment by experienced psychologist **Who diagnosed:** Clinical psychologist **Diagnoses, n** **Among those screening positive** ASD: 7 (high risk) **Non-ASD diagnoses, n** Learning disability: 2 Low risk autism: 3 **Among those screening negative:** NR **Age at diagnosis, mean ± SD:** NR		

Study Description Author/Year: Dietz et al, 2006[16]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: Netherlands **Study period:** October 1999 to April 2002 **Inclusion criteria:** Children screened at well-baby visit at age 14-15 months	**Population characteristics** **Population description:** Children screened at well-baby visit at age 14-15 months **Race/ethnicity, n (%):** NR **Age at screening, mean months ± SD:** 14.91 ± 1.37 **Sex, %:**	**Screening** **Instrument/method:** Children initially prescreened with 4-item ESAT by their doctor; parents of those who screened positive were invited to have a 1.5-hr home visit with a child psychologist who conducted the 14-item ESAT	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **Language development, mean ± SD:** NR	**Performance characteristics: ESAT (%)** **PPV:** 25 **Other outcomes** **Intermediate/Process outcomes** **Time between screening and diagnosis, mean ± SD:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Dietz et al, 2006[16]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Exclusion criteria: • Parents did not cooperate (n=52) • Screening lists incomplete or physician unable to prescreen child, usually due to language barrier (n=399) Note: Population may overlap with Dietz et al[15]	Female: 37% Male: 63% **Risk factors:** NR **IQ, n:** Mullen Scales of Early Learning Low (<70) Screen Positive: 26 Screen Negative: 5 Below average (70-85) Screen Positive: 30 Screen Negative: 22 Average (>85) Screen Positive: 44 Screen Negative: 128 **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD: NR** **Family context:** NR	**Mode of administration:** In-person **Setting:** Clinic; home **Administrator:** Trained child psychologist **Scorer:** Administrator **Study definition of positive screen:** Prescreen: negative answer by parent on 1/4 ESAT items; Screening: ≥3 negative answers in 14-item ESAT by both parent and child psychologist **Total number available to prescreen:** 31,724 **Total number prescreened:** 31,273 **Total number available to screen (test positives from prescreen):** 370 **Total number screened:** 255 (69%) **Number screening positive:** 100 **Number screening negative:** 155 **Percent completing screening followup/diagnosis**: Screen positive: 73 (29% of screened population) Screen negative: None **Diagnosis** **Diagnostic process:** 5 tests over 5 weeks including observation of child behavior in small group of very young children and their parents, standardized parent interview, developmental history, Vineland Social-Emotional Early Childhood Scales, standardized behavior observation using ADOS or ADOS-G, pediatric examination and medical workup **Who diagnosed:** Experienced child psychiatrist made a clinical	**Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ± SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR **Modifiers of screening outcomes/performance:** NR

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Dietz et al, 2006[16]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		judgment about whether the child would meet DSM-IV criteria of different diagnoses by age 3.5 years **Diagnoses, n** **ASD:** 18 **Non-ASD diagnoses, n** **Mental retardation:** 13 **Language disorder:** 22 **Other DSM-IV:** 7 **Other developmental disorder/classification:** 13 **Age at diagnosis, mean ± SD:** NR		

Study Description Author/Year: Baird et al, 2000[18]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
Country: UK **Study period:** NR **Inclusion/exclusion criteria:** Children with profound developmental delay excluded	**Population characteristics** **Population description:** 12-month birth cohort of children in the South East Thames region **Race/ethnicity, n (%)** Caucasian ASD: 101 (100) Other: 95 (100) Typically Developing: 117 (100) **Age at screening, mean ± SD:** 18.7 months ± 1.1 months **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ±SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	**Screening** **Instrument/method:** CHAT **Mode of administration:** Stage 1: Mail (n=2541) and in-person (n=13,694) Stage 2: re-administration of the CHAT 1 month later **Setting:** Primary care clinic **Administrator:** Stage 1: primary health care providers (n=13,694), and via mail (n=2,541) Stage 2: research team **Scorer:** Stage 1: primary health care provider Stage 2: research team **Study definition of positive screen:** High Risk: Failure of all 5 critical items Medium Risk: Failure of item A7 and Biv or only A7 **Total number available to**	**Screening** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** NR **Risk factors:** NR **IQ, mean ±SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR **Diagnosis** **Race/ethnicity, n (%):** NR **Age, mean ± SD:** NR **Sex, n (%):** Female ASD: 3 (3) Aspergers: 0 (0) PDDNOS: 8(9) Male ASD: 42 (45) Aspergers 5 (5) PDDNOS: 36 (38) **Risk factors:** NR	**Performance characteristics (%)** Stage 1 screening, high risk threshold *ASD (including Asperger's)* Sensitivity: 20 Specificity: 99.8 PPV: 26.3 *PDD* Sensitivity: 2.3 Specificity: 99.8 PPV: 2.6 ***All PDDs*** Sensitivity: 11.7 Specificity: 99.8 PPV: 28.9 Stage 1 screening, combined high and medium risk *ASD* Sensitivity: 38 Specificity: 97.6 PPV: 4.7 *PDD* Sensitivity: 31.8

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Baird et al, 2000[18]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		screen: 40,818 **Total number screened:** Stage 1: 16,235 Stage 2: 60 **Number screening positive:** CHAT–Stage 1 High Risk: 38 Medium Risk: 369 CHAT–Stage 2 High Risk: 12 Medium Risk: 22 **Number screening negative:** Stage 1: 15,828 Stage 2: 26/60 **Percent completing screening followup/diagnosis:** CHAT-1: 100 CHAT-2: 0.37 (60 / 16235) **Diagnosis** **Diagnostic process:** Autism Diagnostic Interview–Revised, Leiter International Performance Scale *or* Kaufman Assessment Battery for Children, Reynell Developmental Language Scales *or* Clinical Evaluation of Language Fundamentals–Preschool *or* Clinical Evaluation of Language Fundamentals–Revised; structured interactive assessment **Who diagnosed:** 3 clinicians experienced in diagnosing autism and related disorders **Diagnoses, n** **CHAT-1:** Among screen positives: Autism: 19 PDD: 14 Other diagnosis: 44 Clinically normal: 330 Among screen negatives: Autism: 31	**IQ, mean ±SD:** NR **ASD severity, mean ± SD:** NR **Language development, mean ± SD:** NR **Adaptive behavior, mean ± SD:** NR **Family context:** NR	Specificity: 97.6 PPV: 3.4 *ALL PDDS* Sensitivity: 35.1 Specificity: 97.7 PPV: 8.1 Stage 2 screening**, combined high and medium risk ALL PDDS PPV: 58.8 **Other outcomes*** **Comparison of TP and FN at Initial Screen** ***Gender (n)*** **ASD** TP Female: 2 TP Male: 11 FN Female: 1 FN Male: 12 **PDD** TP Female: 1 TP Male: 19 FN Female: 1 FN Male: 5 **Age at Diagnosis mean ± SD** ***ASD:*** TP=13, FN=13 **TP:** 47.5 ± 14.3 months **FN:** 61.5 ± 12.2 months ***PDD:*** TP=11, FN=6 **TP:** 46.4 ± 7.5 months **FN:** 65 ± 13.8 months **IQ mean ± SD** ***ASD*** **TP:** 65.7 ± 21.8 **FN:** 88.7 ± 16.9 ***PDD*** **TP:** 99.5 ± 15 **FN:** 83.2 ± 10.9 **Intermediate/Process outcomes**

Appendix C. Evidence Table: Screening Studies in Primary Care Populations

Study Description Author/Year: Baird et al, 2000[18]	Baseline/Prescreening Characteristics	Screening and Diagnosis	Characteristics of Population Screened and/or Diagnosed	Screening Outcomes
		PDD: 30 No risk: 15767 CHAT-2: Among screen positives: Autism: 10 PDD: 10 Other diagnosis: 11 Clinically normal: 3 Among screen negatives: NR **Age at diagnosis, mean ± SD:** NR		**Time between screening and diagnosis, mean ± SD:** NR **Time between screening and intervention, mean ± SD:** NR **Health-related outcomes:** NR

Abbreviations: ADHD=Attention Deficit Hyperactivity Disorder; ADI-R=Autism Diagnostic Interview-Revised; ADOS=Autism Diagnostic Observation Schedule; ADOS-G=Autism Diagnostic Observation Schedule-Generic; ADOS-T=Autism Diagnostic Observation Schedule–Toddler Module; ASD=Autism Spectrum Disorder; CARS=Childhood Autism Rating Scale; CHAT=Checklist for Autism in Toddlers; CSBS-DP=Communication and Symbolic Behavior Scale–Developmental Profile; CSBS-DP-T=Communication and Symbolic Behavior Scales–Developmental Profile-Toddlers; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition; ESAT=Early Screening for Autistic Traits; IQ=Intelligence Quotient; ITC=Infant Toddler Checklist; JA-OBS=Joint Attention Observation Schedule; M-CHAT=Modified Checklist for Autism in Toddlers; M-CHAT-F=Modified Checklist for Autism in Toddlers-Follow-up Interview; MSEL=Mullen Scales of Early Learning; NR=Not Reported; PDD-NOS=Pervasive Developmental Disorder- Not Otherwise Specified; SACS=Social Attention and Communication Study; SCQ=Social Communication Questionnaire; SD=Standard Deviation; SES=Socioeconomic Status; VABS=Vineland Adaptive Behavior Scales; YACHTYoung Autism and other developmental disorders Checkup Tool.

References

1. Chlebowski C, Robins DL, Barton ML, et al. Large-scale use of the modified checklist for autism in low-risk toddlers. Pediatrics 2013 Apr;131(4):e1121-7. PMID: 23530174.
2. Robins DL. Screening for autism spectrum disorders in primary care settings. Autism 2008 Sep;12(5):537-56. PMID: 18805946.
3. Kleinman JM, Robins DL, Ventola PE, et al. The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. J Autism Dev Disord 2008 May;38(5):827-39. PMID: 17882539.
4. Pandey J, Verbalis A, Robins DL, et al. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. Autism 2008 Sep;12(5):513-35. PMID: 18805945.
5. Robins DL, Fein D, Barton ML, et al. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. J Autism Dev Disord 2001 Apr;31(2):131-44. PMID: 11450812.
6. Nygren G, Sandberg E, Gillstedt F, et al. A new screening programme for autism in a general population of Swedish toddlers. Res Dev Disabil 2012 Jul-Aug;33(4):1200-10. PMID: 22502846.
7. Barbaro J, Ridgway L, Dissanayake C. Developmental surveillance of infants and toddlers by maternal and child health nurses in an Australian community-based setting: promoting the early identification of autism spectrum disorders. J Pediatr Nurs 2011 Aug;26(4):334-47. PMID: 21726784.
8. Barbaro J, Dissanayake C. Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: the social attention and communication study. J Dev Behav Pediatr 2010 Jun;31(5):376-85. PMID: 20495475.
9. Canal-Bedia R, Garcia-Primo P, Martin-Cilleros MV, et al. Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. J Autism Dev Disord 2011 Oct;41(10):1342-51. PMID: 21161677.
10. Inada N, Koyama T, Inokuchi E, et al. Reliability and validity of the Japanese version of the Modified Checklist for autism in toddlers (M-CHAT). Research in Autism Spectrum Disorders 2011;5(1):330-6. PMID: 815571681; 2010-22960-034.
11. Miller JS, Gabrielsen T, Villalobos M, et al. The each child study: systematic screening for autism spectrum disorders in a pediatric setting. Pediatrics 2011 May;127(5):866-71. PMID: 21482605.

12. Pierce K, Carter C, Weinfeld M, et al. Detecting, studying, and treating autism early: the one-year well-baby check-up approach. J Pediatr 2011 Sep;159(3):458-65 e1-6. PMID: 21524759.
13. Honda H, Shimizu Y, Nitto Y, et al. Extraction and Refinement Strategy for detection of autism in 18-month-olds: a guarantee of higher sensitivity and specificity in the process of mass screening. J Child Psychol Psychiatry 2009 Aug;50(8):972-81. PMID: 19298465.
14. Wetherby AM, Brosnan-Maddox S, Peace V, et al. Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. Autism 2008;12(5):487-511. PMID: 621552511; 2008-13622-005.
15. Dietz C, Swinkels SH, van Daalen E, et al. Parental compliance after screening social development in toddlers. Arch Pediatr Adolesc Med 2007 Apr;161(4):363-8. PMID: 17404133.
16. Dietz C, Swinkels S, van Daalen E, et al. Screening for autistic spectrum disorder in children aged 14-15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. J Autism Dev Disord 2006 Aug;36(6):713-22. PMID: 16633887.
17. VanDenHeuvel A, Fitzgerald M, Greiner B, et al. Screening for autistic spectrum disorder at the 18-month developmental assessment: a population-based study. Ir Med J 2007 Sep;100(8):565-7. PMID: 17955715.
18. Baird G, Charman T, Baron-Cohen S, et al. A screening instrument for autism at 18 months of age: a 6-year follow-up study. J Am Acad Child Adolesc Psychiatry 2000 Jun;39(6):694-702. PMID: 10846303.

Appendix D Table 1. Quality/Internal Validity Ratings for Diagnostic Accuracy Studies

Study, Year	Screening test clearly described	Repre-sentative population	Selection criteria clearly described	Credible reference standard	Reference standard time period appropriate	Reference test applied to all or subset of participants	Same reference standard applied to all participants	Reference standard independ-ent of screening test	Replication of screening test possible	Replication of reference standard possible	Screening test and reference standard interpreted independ-ently	Same data used to interpret test available if test used in clinical practice	With-drawals from study explained	Methods/ data for calcul-ating accuracy clearly reported	Methods/ data for calcul-ating accuracy valid	Internal validity rating
Chlebowski 2013[1-5]	yes	yes	yes	yes	yes	all	yes	yes	yes	yes	no	yes	yes	yes	yes	Good
Nygren 2012[6]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	unclear	yes	yes	yes	yes	Fair
Barbaro 2011[7,8]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	no	yes	yes	yes	yes	Fair
Canal-Bedia 2011a[9]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	yes	yes	no	yes	yes	Fair
Canal-Bedia 2011b[9]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	yes	yes	no	yes	yes	Fair
Inada 2011[10]	yes	yes	yes	yes	yes	all	yes	yes	yes	yes	yes	yes	yes	yes	yes	Fair-Poor
Miller 2011[11]	yes	yes	yes	yes	yes	subset	no	yes	yes	yes	yes	yes	yes	yes	yes	Fair
Pierce 2011[12]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	no	yes	yes	yes	yes	Good
Honda 2009[13]	yes	unclear	yes	yes	yes	subset	yes	yes	yes	yes	yes	yes	yes	yes	yes	Fair
Wetherby 2008[14]	yes	yes	yes	yes	yes	all	yes	yes	yes	yes	yes	yes	yes	yes	yes	Good
Oosterling 2007[15]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	yes	yes	no	yes	yes	Fair
Vanden Heuvel 2007[16]	yes	yes	yes	unclear	yes	subset	no	no	no	no	unclear	unclea r	yes	no	yes	Poor
Oosterling 2006[17]	yes	yes	yes	yes	yes	subset	yes	yes	yes	yes	yes	yes	no	yes	yes	Fair
Baird 2000[18]	yes	yes	yes	yes	yes	all	yes	yes	yes	yes	yes	yes	yes	yes	yes	Good

References

1. Kleinman JM, Robins DL, Ventola PE, et al. The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. J Autism Dev Disord 2008 May;38(5):827-39. PMID: 17882539.
2. Chlebowski C, Robins DL, Barton ML, et al. Large-scale use of the modified checklist for autism in low-risk toddlers. Pediatrics 2013 Apr;131(4):e1121-7. PMID: 23530174.
3. Robins DL. Screening for autism spectrum disorders in primary care settings. Autism 2008 Sep;12(5):537-56. PMID: 18805946.
4. Pandey J, Verbalis A, Robins DL, et al. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. Autism 2008 Sep;12(5):513-35. PMID: 18805945.

5. Robins DL, Fein D, Barton ML, et al. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. J Autism Dev Disord 2001 Apr;31(2):131-44. PMID: 11450812.
6. Nygren G, Sandberg E, Gillstedt F, et al. A new screening programme for autism in a general population of Swedish toddlers. Res Dev Disabil 2012 Jul-Aug;33(4):1200-10. PMID: 22502846.
7. Heyer NJ, Echeverria D, Woods JS. Disordered porphyrin metabolism: a potential biological marker for autism risk assessment. Autism Res 2012 Apr;5(2):84-92. PMID: 22298513.
8. Barbaro J, Dissanayake C. Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: the social attention and communication study. J Dev Behav Pediatr 2010 Jun;31(5):376-85. PMID: 20495475.
9. Canal-Bedia R, Garcia-Primo P, Martin-Cilleros MV, et al. Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. J Autism Dev Disord 2011 Oct;41(10):1342-51. PMID: 21161677.
10. Inada N, Koyama T, Inokuchi E, et al. Reliability and validity of the Japanese version of the Modified Checklist for autism in toddlers (M-CHAT). Research in Autism Spectrum Disorders 2011;5(1):330-6. PMID: 815571681; 2010-22960-034.
11. Miller JS, Gabrielsen T, Villalobos M, et al. The each child study: systematic screening for autism spectrum disorders in a pediatric setting. Pediatrics 2011 May;127(5):866-71. PMID: 21482605.
12. Pierce K, Carter C, Weinfeld M, et al. Detecting, studying, and treating autism early: the one-year well-baby check-up approach. J Pediatr 2011 Sep;159(3):458-65 e1-6. PMID: 21524759.
13. Honda H, Shimizu Y, Nitto Y, et al. Extraction and Refinement Strategy for detection of autism in 18-month-olds: a guarantee of higher sensitivity and specificity in the process of mass screening. J Child Psychol Psychiatry 2009 Aug;50(8):972-81. PMID: 19298465.
14. Wetherby AM, Brosnan-Maddox S, Peace V, et al. Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. Autism 2008;12(5):487-511. PMID: 621552511; 2008-13622-005.
15. Dietz C, Swinkels SH, van Daalen E, et al. Parental compliance after screening social development in toddlers. Arch Pediatr Adolesc Med 2007 Apr;161(4):363-8. PMID: 17404133.
16. VanDenHeuvel A, Fitzgerald M, Greiner B, et al. Screening for autistic spectrum disorder at the 18-month developmental assessment: a population-based study. Ir Med J 2007 Sep;100(8):565-7. PMID: 17955715.
17. Dietz C, Swinkels S, van Daalen E, et al. Screening for autistic spectrum disorder in children aged 14-15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. J Autism Dev Disord 2006 Aug;36(6):713-22. PMID: 16633887.
18. Baird G, Charman T, Baron-Cohen S, et al. A screening instrument for autism at 18 months of age: a 6-year follow-up study. J Am Acad Child Adolesc Psychiatry 2000 Jun;39(6):694-702. PMID: 10846303.

Appendix D Table 2. Quality Assessment Form for Diagnostic Accuracy Studies

Screening for ASD in Young Children
Quality scoring—Diagnostic accuracy studies

Reviewer: _____ RefID:_____

Question (Considerations)	Circle your assessment	Comments/Rationale
1. Test(s) clearly described (or referenced)?	Yes No	
2. Was the spectrum of patients representative of the patients who will receive the test in primary care? (For studies with children with impairments, consider if likely to be representative of the larger impaired population or does sample differ markedly in some way)	Yes No	
3. Were selection criteria clearly described?	Yes No	
4. Is the reference standard likely to correctly classify the target condition? (Diagnosis should include one of the following: A. clinical DSM-IV-based diagnosis + ADI-R + ADOS B. clinical DSM-IV-based diagnosis + ADOS OR ADI-R C. [clinical DSM-IV-based diagnosis + other] OR [ADOS + other, such as SRS, CARS, SCQ, CAST, ASSQ, OR STAT, MCHAT for under 30 months] D. Only clinical DSM-IV-based diagnosis OR Only ADOS)	Yes No	
5. Is the time period of the reference standard appropriate to confirm diagnosis? (Diagnosis generally between 2-4 years of age.)	Yes No	
6. Did the whole or random/selected sample receive reference test?	Whole Selected	
7. Did patients receive the same reference regardless of test results?	Yes No	
8. Was the reference standard independent of the test?	Yes No	
9. Was the execution test described in enough details to permit replication of the test?	Yes No	
10. Was the execution of the reference standard described in enough detail to permit replication of the test?	Yes No	
11. Were the index test and reference standard results interpreted independently (blinded)?	Yes No	
12. Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?	Yes No	
13. Were withdrawals from the study explained (post enrollment)?	Yes No	
14. Methods/data for calculating accuracy clearly reported?	Yes No	
15. Methods/data for calculating accuracy valid?	Yes No	

Appendix D Table 3. Quality Ratings for Intervention Studies

First Author Year	Group Design	Random Assign-ment	Appropriate Comparison Group	Correct Random-ization	Systematic Diagnostic Approach	Clear Sample Character-ization	Clear Inclusion/ Exclusion Criteria	Attrition Reported	Dropout Character-istics Evaluated	Inter-vention Fully Described	Treatment Fidelity Monitored	Concomitant Interventions Held Steady/ Reported	Outcome Measures Reliable and Valid	Primary Outcomes Specified *a priori*	Outcome Data Collected From Appropriate Sources	Out-comes Coded Blindly	Appro-priate Statistical Analysis	Rating
Kasari 2014[1]	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	F
Boyd 2013[2]	+	-	+	NA	+	+	+	+	+	+	+	-	+	+	+	-	-	F
Casenhiser 2013[3]	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	F
Goods 2013[4]	+	+	+	+	+	+	+	+	+	+	+	NA	+	+	+	+	+	G
Peters-Scheffer 2013[5]	+	-	+	NA	+	+	+	+	-	+	+	-	+	-	+	-	+	G
Schertz 2013[6]	+	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	+	F
Schreibman 2013 [7]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	G
Siller 2013[8]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F
Warreyn 2013[9]	+	+	+	-	+	+	+	+	-	+	-	-	+	+	+	-	-	F
Wong 2013[10]	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	+	F
Dawson 2012[11, 12]	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	G
Eikeseth 2012[13]	+	-	+	NA	-	+	-	+	-	+	-	+	+	+	+	-	-	F
Eldevik 2012[14]	+	-	+	NA	-	+	+	+	+	+	-	-	+	-	+	-	+	F
Flanagan 2012[15, 16]	+	-	+	NA	+	+	+	NA	-	+	-	-	+	+	+	-	+	F
Kaale 2012[17]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F
Landa 2012[18, 19]	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	G
Lawton 2012[20]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	F
Reed 2012[21]	+	-	+	NA	+	+	+	-	-	+	-	-	+	+	+	+	-	F
Rogers 2012[22, 23]	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	-	-	F
Strauss 2012[24, 25]	+	-	+	NA	+	+	+	+	+	+	+	-	+	+	+	-	+	G

Appendix D Table 3. Quality Ratings for Intervention Studies

First Author Year	Group Design	Random Assign-ment	Appropriate Comparison Group	Correct Random-ization	Systematic Diagnostic Approach	Clear Sample Character-ization	Clear Inclusion/ Exclusion Criteria	Attrition Reported	Dropout Character-istics Evaluated	Inter-vention Fully Described	Treatment Fidelity Monitored	Concomitant Interventions Held Steady/ Reported	Outcome Measures Reliable and Valid	Primary Outcomes Specified *a priori*	Outcome Data Collected From Appropriate Sources	Out-comes Coded Blindly	Appro-priate Statistical Analysis	Rating
Venker 2012[26]	+	+	+	-	+	+	-	-	-	+	+	-	+	+	+	-	+	F
Aldred 2011[27, 28]	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	G
Carter 2011[29]	+	+	+	+	+	+	+	+	-	+	+	-	+	+	+	+	+	F
Itzchak 2011[30, 31]	+	-	+	NA	+	+	+	-	-	+	-	-	+	+	+	-	-	F
Landa 2011[18, 19]	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	G
Roberts 2011[32]	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	G
Strain 2011[33]	+	+	+	-	-	+	+	+	+	+	+	-	+	+	+	-	+	F
Green 2010[34]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F
Ingersoll 2010[35, 36]	+	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	+	G
Itzchak 2010[30, 31]	+	-	+	NA	+	+	+	-	-	+	-	-	+	+	+	-	-	F
Kasari 2010[37]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	G
Keen 2010[38]	+	-	+	NA	+	+	-	+	+	+	+	-	+	+	+	-	+	G
Oosterling 2010[39]	+	+	+	-	+	+	+	+	-	+	+	-	+	+	+	-	-	F
Peters-Sheffer 2010[40]	+	-	+	NA	+	+	+	+	+	+	-	+	+	+	+	-	+	F
Reed 2010[21]	+	-	+	NA	+	+	+	-	-	+	-	-	+	+	+	+	-	F
Siller 2010[8]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F
Hayward 2009[41, 42]	+	-	+	NA	-	+	+	+	+	+	+	-	+	+	+	+	+	F
Reed 2007[43]	+	-	+	NA	-	+	+	+	-	+	-	-	+	+	+	-	+	F
Zachor 2007[44]	+	-	+	NA	+	+	-	+	N	+	-	-	+	+	+	-	+	F
Cohen 2006[45]	+	-	+	NA	+	+	+	+	+	+	-	-	+	+	+	-	+	F

Appendix D Table 3. Quality Ratings for Intervention Studies

First Author Year	Group Design	Random Assign-ment	Appropriate Comparison Group	Correct Random-ization	Systematic Diagnostic Approach	Clear Sample Character-ization	Clear Inclusion/ Exclusion Criteria	Attrition Reported	Dropout Character-istics Evaluated	Inter-vention Fully Described	Treatment Fidelity Monitored	Concomitant Interventions Held Steady/ Reported	Outcome Measures Reliable and Valid	Primary Outcomes Specified *a priori*	Outcome Data Collected From Appropriate Sources	Out-comes Coded Blindly	Appro-priate Statistical Analysis	Rating
Kasari 2006[46-49]	+	+	+	-	+	+	+	+	-	+	+	-	+	+	+	+	+	F
Howard 2005[50]	+	-	+	N	+	+	+	+	-	+	-	-	+	+	+	-	+	F
Sallows 2005[51]	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	G
Aldred 2004[27, 28]	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	G
Drew 2002[52]	+	+	+	+	-	+	+	+	+	+	-	+	+	+	+	-	+	F
Eikeseth 2002[53, 54]	+	-	+	NA	-	+	+	+	-	+	-	-	+	+	+	+	+	F
Smith 2000[21]	+	+	+	-	-	+	+	+	+	+	+	-	+	+	+	+	+	F

Abbreviations: F=fair; G=good; NA=not applicable.

References

1. Kasari C, Lawton K, Shih W, et al. Caregiver-mediated intervention for low-resourced preschoolers with autism: an RCT. Pediatrics 2014 Jul;134(1):e72-9. PMID: 24958585.
2. Boyd BA, Hume K, McBee MT, et al. Comparative Efficacy of LEAP, TEACCH and Non-Model-Specific Special Education Programs for Preschoolers with Autism Spectrum Disorders. J Autism Dev Disord 2013 Jun 28PMID: 23812661.
3. Casenhiser DM, Shanker SG, Stieben J. Learning through interaction in children with autism: preliminary data from asocial-communication-based intervention. Autism 2013 Mar;17(2):220-41. PMID: 21949005.
4. Goods KS, Ishijima E, Chang Y-C, et al. Preschool Based JASPER Intervention in Minimally Verbal Children with Autism: Pilot RCT. Journal of Autism and Developmental Disorders 2013 May 2013;43(5):1050-6. PMID: 1361838997; EJ1000575.
5. Peters-Scheffer N, Didden R, Mulders M, et al. Effectiveness of low intensity behavioral treatment for children with autism spectrum disorder and intellectual disability. Research in Autism Spectrum Disorders 2013;7(9):1012-25. PMID: 1468570233; 2013-27709-003.
6. Schertz HH, Odom SL, Baggett KM, et al. Effects of Joint Attention Mediated Learning for Toddlers with Autism Spectrum Disorders: An Initial Randomized Controlled Study. Early Childhood Research Quarterly 2013 2013;28(2):249-58. PMID: 1413414144; EJ1008279.
7. Schreibman L, Stahmer AC. A Randomized Trial Comparison of the Effects of Verbal and Pictorial Naturalistic Communication Strategies on Spoken Language for Young Children with Autism. J Autism Dev Disord 2013 Nov 23PMID: 24272416.
8. Siller M, Hutman T, Sigman M. A Parent-Mediated Intervention to Increase Responsive Parental Behaviors and Child Communication in Children with ASD: A Randomized Clinical Trial. Journal of Autism and Developmental Disorders 2013;43(3):540-55. PMID: 1347458326; EJ999150.
9. Warreyn P, Roeyers H. See what I see, do as I do: Promoting joint attention and imitation in preschoolers with autism spectrum disorder. Autism 2013 Oct 8PMID: 24104513.
10. Wong CS. A play and joint attention intervention for teachers of young children with Autism: A randomized controlled pilot study. Autism 2013 May 2013;17(3):340-57. PMID: 1463475195; 18038415.
11. Dawson G, Rogers S, Munson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics 2010 Jan;125(1):e17-23. PMID: 19948568.

12. Dawson G, Jones EJ, Merkle K, et al. Early behavioral intervention is associated with normalized brain activity in young children with autism. J Am Acad Child Adolesc Psychiatry 2012 Nov;51(11):1150-9. PMID: 23101741.
13. Eikeseth S, Klintwall L, Jahr E, et al. Outcome for Children with Autism Receiving Early and Intensive Behavioral Intervention in Mainstream Preschool and Kindergarten Settings. Research in Autism Spectrum Disorders 2012;6(2):829-35. PMID: 1023529557; EJ966831.
14. Eldevik S, Hastings RP, Jahr E, et al. Outcomes of behavioral intervention for children with autism in mainstream pre-school settings. J Autism Dev Disord 2012 Feb;42(2):210-20. PMID: 21472360.
15. Flanagan HE, Perry A, Freeman NL. Effectiveness of large-scale community-based intensive Behavioral Intervention: A waitlist comparison study exploring outcomes and predictors. Research in Autism Spectrum Disorders 2012;6(2):673-82. PMID: 928983882; 2012-03990-011.
16. Perry A, Blacklock K, Dunn Geier J. The relative importance of age and IQ as predictors of outcomes in Intensive Behavioral Intervention. Research in Autism Spectrum Disorders 2013;7(9):1142-50. PMID: 1468570295; 2013-27709-017.
17. Kaale A, Smith L, Sponheim E. A randomized controlled trial of preschool-based joint attention intervention for children with autism. J Child Psychol Psychiatry 2012 Jan;53(1):97-105. PMID: 21883204.
18. Landa RJ, Holman KC, O'Neill AH, et al. Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: a randomized controlled trial. J Child Psychol Psychiatry 2011 Jan;52(1):13-21. PMID: 21126245.
19. Landa RJ, Kalb LG. Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention. Pediatrics 2012 Nov;130 Suppl 2:S186-90. PMID: 23118250.
20. Lawton K, Kasari C. Teacher-implemented joint attention intervention: pilot randomized controlled study for preschoolers with autism. J Consult Clin Psychol 2012 Aug;80(4):687-93. PMID: 22582764.
21. Reed P, Osborne L. Impact of Severity of Autism and Intervention Time-Input on Child Outcomes: Comparison across Several Early Interventions. British Journal of Special Education 2012;39(3):130-6. PMID: 1140138033; EJ979673.
22. Rogers SJ, Estes A, Lord C, et al. Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. Journal of the American Academy of Child & Adolescent Psychiatry 2012;51(10):1052-65.
23. Estes A, Vismara L, Mercado C, et al. The Impact of Parent-Delivered Intervention on Parents of Very Young Children with Autism. J Autism Dev Disord 2013 Jul 10PMID: 23838727.
24. Strauss K, Vicari S, Valeri G, et al. Parent inclusion in Early Intensive Behavioral Intervention: the influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. Res Dev Disabil 2012 Mar-Apr;33(2):688-703. PMID: 22188793.
25. Fava L, Strauss K, Valeri G, et al. The Effectiveness of a Cross-Setting Complementary Staff- and Parent-Mediated Early Intensive Behavioral Intervention for Young Children with ASD. Research in Autism Spectrum Disorders 2011;5(4):1479-92. PMID: 881460693; EJ927555.
26. Venker CE, McDuffie A, Weismer SE, et al. Increasing Verbal Responsiveness in Parents of Children with Autism: A Pilot Study. Autism: The International Journal of Research and Practice 2012;16(6):568-85. PMID: 1238188272; EJ984594.
27. Aldred C, Green J, Adams C. A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness. J Child Psychol Psychiatry 2004 Nov;45(8):1420-30. PMID: 15482502.
28. Aldred C, Green J, Emsley R, et al. Mediation of treatment effect in a communication intervention for pre-school children with autism. Journal of Autism and Developmental Disorders 2012;42(3):447-54. PMID: 928983846; 2012-04398-013.
29. Carter AS, Messinger DS, Stone WL, et al. A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. Journal of Child Psychology and Psychiatry 2011;52(7):741-52.
30. Itzchak EB, Zachor DA. Who Benefits from Early Intervention in Autism Spectrum Disorders? Research in Autism Spectrum Disorders 2011;5(1):345-50. PMID: 762466666; EJ900428.
31. Zachor DA, Itzchak EB. Treatment Approach, Autism Severity and Intervention Outcomes in Young Children. Research in Autism Spectrum Disorders 2010;4(3):425-32. PMID: 742864417; EJ878590.
32. Roberts J, Williams K, Carter M, et al. A Randomised Controlled Trial of Two Early Intervention Programs for Young Children with Autism: Centre-Based with Parent Program and Home-Based. Research in Autism Spectrum Disorders 2011;5(4):1553-66. PMID: 881469872; EJ927523.
33. Strain PS, Bovey EH. Randomized, Controlled Trial of the Leap Model of Early Intervention for Young Children with Autism Spectrum Disorders. Topics in Early Childhood Special Education 2011;31(3):133-54. PMID: 964193023; EJ942105.

34. Green J, Charman T, McConachie H, et al. Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. Lancet 2010 May 20PMID: 20494434.
35. Ingersoll B. Brief report: Pilot randomized controlled trial of reciprocal imitation training for teaching elicited and spontaneous imitation to children with autism. Journal of Autism and Developmental Disorders 2010;40(9):1154-60. PMID: 754055593; 2010-17107-012.
36. Ingersoll B. Brief report: effect of a focused imitation intervention on social functioning in children with autism. J Autism Dev Disord 2012 Aug;42(8):1768-73. PMID: 22146934.
37. Kasari C, Gulsrud AC, Wong C, et al. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. Journal of Autism and Developmental Disorders 2010;40(9):1045-56. PMID: 754055821; 2010-17107-001.
38. Keen D, Couzens D, Muspratt S, et al. The effects of a parent-focused intervention for children with a recent diagnosis of autism spectrum disorder on parenting stress and competence. Research in Autism Spectrum Disorders 2010;4(2):229-41. PMID: 856409394; 2011-03277-009.
39. Oosterling I, Visser J, Swinkels S, et al. Randomized controlled trial of the focus parent training for toddlers with autism: 1-year outcome. J Autism Dev Disord 2010 Dec;40(12):1447-58. PMID: 20440639.
40. Peters-Scheffer N, Didden R, Mulders M, et al. Low intensity behavioral treatment supplementing preschool services for young children with autism spectrum disorders and severe to mild intellectual disability. Res Dev Disabil 2010 Nov-Dec;31(6):1678-84. PMID: 20627451.
41. Hayward D, Eikeseth S, Gale C, et al. Assessing progress during treatment for young children with autism receiving intensive behavioural interventions. Autism 2009 Nov;13(6):613-33. PMID: 19933766.
42. Eikeseth S, Hayward D, Gale C, et al. Intensity of Supervision and Outcome for Preschool Aged Children Receiving Early and Intensive Behavioral Interventions: A Preliminary Study. Research in Autism Spectrum Disorders 2009 Jan;3(1):67-73. PMID: EJ820616.
43. Reed P, Osborne LA, Corness M. Brief report: relative effectiveness of different home-based behavioral approaches to early teaching intervention. J Autism Dev Disord 2007 Oct;37(9):1815-21. PMID: 17180714.
44. Zachor DA, Ben-Itzchak E, Rabinovich A-L, et al. Change in autism core symptoms with intervention. Research in Autism Spectrum Disorders 2007;1(4):304-17.
45. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S145-55. PMID: 16685181.
46. Kasari C, Freeman S, Paparella T. Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. J Child Psychol Psychiatry 2006 Jun;47(6):611-20. PMID: 16712638.
47. Kasari C, Gulsrud A, Freeman S, et al. Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. J Am Acad Child Adolesc Psychiatry 2012 May;51(5):487-95. PMID: 22525955.
48. Kasari C, Paparella T, Freeman S, et al. Language outcome in autism: randomized comparison of joint attention and play interventions. J Consult Clin Psychol 2008 Feb;76(1):125-37. PMID: 18229990.
49. Lawton K, Kasari C. Brief Report: Longitudinal Improvements in the Quality of Joint Attention in Preschool Children with Autism. Journal of Autism and Developmental Disorders 2012;42(2):307-12. PMID: 964190966; EJ954567.
50. Howard JS, Sparkman CR, Cohen HG, et al. A comparison of intensive behavior analytic and eclectic treatments for young children with autism. Res Dev Disabil 2005 Jul-Aug;26(4):359-83. PMID: 15766629.
51. Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four-year outcome and predictors. Am J Ment Retard 2005 Nov;110(6):417-38. PMID: 16212446.
52. Drew A, Baird G, Baron-Cohen S, et al. A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. Eur Child Adolesc Psychiatry 2002 Dec;11(6):266-72. PMID: 12541005.
53. Eikeseth S, Smith T, Jahr E, et al. Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. Behav Modif 2002 Jan;26(1):49-68. PMID: 11799654.
54. Eikeseth S, Smith T, Jahr E, et al. Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: a comparison controlled study. Behav Modif 2007 May;31(3):264-78. PMID: 17438342.

Appendix D Table 4. Quality/Risk of Bias Assessment Approach

Study Design

1. Did the study employ a group design?
   Group designs may include randomized controlled trials, prospective or retrospective cohorts, case-control studies
   + = yes
   - = no
2. Were the groups randomly assigned?
   + = yes
   - = no
3. Was there an appropriate comparison group?
   The comparison group should accurately represent the characteristics of the intervention group in the absence of the intervention. Specifically, factors that are likely to be associated with the intervention selected and with outcomes observed should be evenly distributed between groups, if possible. These factors may include, for example, age, IQ, severity, etc.
   + = yes
   - = no or not reported (NR)
4. If an RCT, was randomization done correctly?
   + = yes
   - = no
   NR
   NA for all non-RCTs
   Considerations:
   Was the approach to randomization described? Were random techniques like computer-generated, sequentially numbered opaque envelope used?
   Were technically non-random techniques, like alternate days of the week used?
   Any studies with randomization techniques not reported (NR) will also be reviewed by the team.

Participant Ascertainment/Inclusion

1. Was a valid diagnostic approach for ASD used within the study, or were referred participants diagnosed using a valid approach?
   A. clinical DSM-IV-based diagnosis + ADI-R and/or ADOS
   B. [clinical DSM-IV-based diagnosis + other] OR [ADOS + other, such as SRS, CARS, SCQ, CAST, ASSQ, OR STAT, MCHAT for under 30 months]
   C. Only clinical DSM-IV-based diagnosis OR Only ADOS
   D. Neither clinical DSM-IV-based diagnosis NOR ADOS
2. Was the sample clearly characterized (e.g., information provided to characterize participants in terms of impairments associated with their ASD, such as cognitive or developmental level)?
   + = yes
   - = no or not reported (NR)
   Considerations:
   Are baseline measures of IQ, mental age, language facility, etc. reported?
   How reproducible is the study in terms of the sample participants? Do the authors provide enough information that you could recreate the study population in a new study?
3. Were inclusion and exclusion criteria clearly stated?
   + = yes
   - = no or not reported (NR)
   Considerations:
   Did the authors report this information?
4. Do the authors report attrition?
   + = yes
   - = no
   Considerations:
   Do they report loss to follow-up and/or drop-out?
   If there is no attrition (i.e., baseline and follow up Ns are the same), score as YES
5. Were characteristics of drop-out group evaluated for differences with the participant group as a whole?
   + = yes

Appendix D Table 4. Quality/Risk of Bias Assessment Approach

- = no or not reported (NR)
NA or attrition was minimal
Considerations:
Were reasons for dropping out evaluated?
Does the paper describe a comparison between drop-outs and the whole group?
Score as NA if attrition was minimal.

Intervention
1. Was the intervention fully described?
+ = yes
- = no or not reported (NR)
Considerations:
Is there sufficient detail to allow replication of the intervention?
Does the study describe the dosage, formulation, timing, duration, intensity, etc. of the intervention?
Do the authors refer to a treatment manual (score as YES if so, even is manual is unpublished)?
2. For behavioral studies, was treatment fidelity monitored in a systematic way?
+ = yes
- = no or not reported (NR)
NA
Considerations:
Was a method in place to assess whether people providing he intervention were adherent to a manual/process? We're not assessing the quality of the fidelity, just whether it was performed.
3. Did the authors measure and report adherence to the intended treatment process?
+ = yes
- = no or not reported (NR)
Considerations:
Does the study report number of hours of treatment or treatment sessions or time period receiving therapy (planned vs. actually received)? Do they provide pill count data or parental medication diary, etc. for pharmacologic interventions?
4. Did the authors report differences in or hold steady all concomitant interventions?
+ = yes
- = no or not reported (NR)
Considerations:
Was an attempt made to assess/determine if other interventions were ongoing?

Outcome Measurement
1. Did outcome measures demonstrate adequate reliability and validity (including interobserver reliability for behavior observation coding)?
+ = yes
- = no or not reported (NR)
Considerations:
If the study used an established measure, has validity been established previously and do the authors provide a reference?
If the study used a new measure, was validity established?
For interobserver coding, was reliability and /or validity tested?
2. Were the primary & secondary outcomes clearly specified a priori?
+ = yes
- = no or not reported (NR)
Considerations:
Was there a "called shot?"
3. Were outcome data collected from sources appropriate to the target outcome (e.g. parent report, teacher report, direct behavior observation)?
+ = yes
- = no or not reported (NR)
Considerations:
Ex: Parent report for home-focused outcomes, teacher report for academic/school-focused, etc.

Appendix D Table 4. Quality/Risk of Bias Assessment Approach

4. Were outcomes coded by individuals blinded to the intervention status of the participants?
   + = yes
   - = no or not reported (NR)

Analysis

1. Was an appropriate statistical analysis used?
   + = yes
   - = no
   a. For RCT's, was there an intent-to treat analysis?
      + = yes
      - = no
      NA
      Considerations:
      Does the study report ITT analyses or last observation carried forward or note that all subjects were included in the final analyses?
      If ≤2 participants were lost to follow-up, consider the analysis as ITT.
   b. For negative studies, was a power calculation provided?
      + = yes
      - = no
      NA
   c. Did the study correct for multiple testing?
      + = yes
      - = no
      NA
   d. For observational studies, were potential confounders and effect measure modifiers captured?
      + = yes
      - = no
      NA
      Considerations:
      Were the groups well categorized at baseline? Were baseline differences assessed?
   e. For observational studies, were potential confounders and effect measure modifiers handled appropriately?
      + = appropriate analysis
      - = inappropriate analysis
      NA
      Considerations:
      Confounders are variables that are associated both with the intervention and the outcome and that change the relationship of the intervention to the outcome. These are variables that we would control for in analysis.
      Effect measure modifiers are variables that we think of as stratifying, in that the relationship between the intervention and outcome is fundamentally different in different strata of the effect modifier. Observational research should include an assessment of potential confounders and modifiers, and if they are observed, analysis should control for or stratify on them.
      Was the candidate variable selection discussed/noted?
      Was the model-building approach described?
      Were any variables unrelated to the studied variables that could have altered the outcome handled appropriately?
      Were any variables not under study that affected the causal factors handled appropriately?

Appendix E. Excluded Studies

Key to Exclusion Reasons
X-1: Did not address screening
X-2: Not original research
X-3: Ineligible population
X-4: Did not address outcomes of interest
X-5: Not conducted in eligible country
X-6: Article not obtainable

1. Baron-Cohen S, Wheelwright S, Cox A, et al. Early identification of autism by the CHecklist for Autism in Toddlers (CHAT). J R Soc Med. 2000 Oct;93(10):521-5. PMID: 11064690; X-2
2. Eaves RC, Campbell HA, Chambers D. Criterion-related and construct validity of the Pervasive Developmental Disorders Rating Scale and the Autism Behavior Checklist. Psychology in the Schools. 2000;37(4):311-21. X-3
3. Fenson L, Pethick S, Renda C, et al. Short-Form Versions of the MacArthur Communicative Development Inventories. Applied Psycholinguistics. 2000;21(1):95-115. X-1, X-3, X-4
4. Stone WL, Coonrod EE, Ousley OY. Brief report: screening tool for autism in two-year-olds (STAT): development and preliminary data. J Autism Dev Disord. 2000 Dec;30(6):607-12. PMID: 11261472; X-4
5. Adrien JL, Rossignol-Deletang N, Martineau J, et al. Regulation of cognitive activity and early communication development in young autistic, mentally retarded, and young normal children. Dev Psychobiol. 2001 Sep;39(2):124-36. PMID: 11568882; X-3, X-4
6. Baird G, Charman T, Cox A, et al. Current topic: Screening and surveillance for autism and pervasive developmental disorders. Arch Dis Child. 2001 Jun;84(6):468-75. PMID: 11369559; X-2, X-3, X-4, X-5
7. Baird G, Charman T, Santosh PJ. Clinical considerations in the diagnosis of autism spectrum disorders. Indian J Pediatr. 2001 May;68(5):439-49. PMID: 11407161; X-1, X-2, X-3, X-4, X-5
8. Blackwell J. Clinical practice guideline: screening and diagnosing autism. J Am Acad Nurse Pract. 2001 Dec;13(12):534-6. PMID: 11836828; X-2, X-3, X-4, X-5
9. Chakrabarti S, Fombonne E. Pervasive developmental disorders in preschool children. JAMA. 2001 Jun 27;285(24):3093-9. PMID: 11427137; X-1, X-4
10. Charman T, Baron-Cohen S, Baird G, et al. Commentary: The Modified Checklist for Autism in Toddlers. Journal of Autism and Developmental Disorders. 2001;31(2):145-8. X-2, X-3, X-4, X-5
11. Robins DL, Fein D, Barton ML, et al. Reply to Charman et al.'s Commentary on the Modified Checklist for Autism in Toddlers. Journal of Autism and Developmental Disorders. 2001;31(2):149-51. X-2, X-3, X-4
12. Scambler D, Rogers SJ, Wehner EA. Can the checklist for autism in toddlers differentiate young children with autism from those with developmental delays? J Am Acad Child Adolesc Psychiatry. 2001 Dec;40(12):1457-63. PMID: 11765292; X-4
13. Willemsen-Swinkels SH, Buitelaar JK, van Engeland H. Is 18 months too early for the CHAT? J Am Acad Child Adolesc Psychiatry. 2001 Jul;40(7):737-8. PMID: 11437006; X-2, X-4
14. Zwaigenbaum L. Autistic spectrum disorders in preschool children. Can Fam Physician. 2001 Oct;47:2037-42. PMID: 11723598; X-2, X-3, X-4, X-5
15. Brereton AV, Tonge BJ, Mackinnon AJ, et al. Screening young people for autism with the developmental behavior checklist. J Am Acad Child Adolesc Psychiatry. 2002 Nov;41(11):1369-75. PMID: 12410080; X-3, X-4
16. Drew A, Baird G, Baron-Cohen S, et al. A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. Eur Child Adolesc Psychiatry. 2002 Dec;11(6):266-72. PMID: 12541005; X-1, X-4
17. Honda H, Shimizu Y. Early Intervention System for Preschool Children with Autism in the Community: The DISCOVERY Approach in Yokohama, Japan. Autism: The International Journal of Research and Practice. 2002;6(3):239-57. X-4
18. Jensen VK, Sinclair LV. Treatment of Autism in Young Children: Behavioral Intervention and Applied Behavior Analysis. Infants and Young Children. 2002;14(4):42-52. X-1, X-2, X-3, X-4
19. Plotts C, Webber J. The Role of Developmental Histories in the Screening and Diagnosis of Autism Spectrum Disorders. Assessment for Effective Intervention. 2002;27(1-2):19-26. X-1, X-2, X-3, X-4
20. South M, Williams BJ, McMahon WM, et al. Utility of the Gilliam Autism Rating Scale in Research and Clinical Populations. Journal of Autism and Developmental Disorders. 2002;32(6):593-9. X-3, X-4
21. Sprafkin J, Volpe RJ, Gadow KD, et al. A DSM-IV-referenced screening instrument for preschool children: the Early Childhood Inventory-4. J Am Acad Child Adolesc Psychiatry. 2002 May;41(5):604-12. PMID: 12014793; X-1, X-3, X-4
22. Vrancic D, Nanclares V, Soares D, et al. Sensitivity and specificity of the autism diagnostic inventory-telephone screening in Spanish. J Autism Dev Disord. 2002 Aug;32(4):313-20. PMID: 12199136; X-3, X-4
23. Williams TO, Eaves RC. The reliability of test scores for the Pervasive Developmental Disorders Rating Scale. Psychology in the Schools. 2002;39(6):605-11. X-4
24. Blackwell J, Niederhauser C. Diagnose and manage autistic children. Nurse Pract. 2003 Jun;28(6):36-43; quiz 4-5. PMID: 12796622; X-2, X-3, X-4, X-5

25. Manning-Courtney P, Brown J, Molloy CA, et al. Diagnosis and treatment of autism spectrum disorders. Curr Probl Pediatr Adolesc Health Care. 2003 Oct;33(9):283-304. PMID: 14534514; X-2, X-3, X-4, X-5
26. Nash PL, Coury DL. Screening tools assist with diagnosis of autistic spectrum disorders. Pediatr Ann. 2003 Oct;32(10):664-70. PMID: 14606216; X-2, X-3, X-4, X-5
27. Saemundsen E, Magnusson P, Smari J, et al. Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: Convergence and Discrepancy in Diagnosing Autism. Journal of Autism and Developmental Disorders. 2003;33(3):319-28. X-4
28. Stone WL, Coonrod EE, Pozdol SL, et al. The Parent Interview for Autism-Clinical Version (PIA-CV): A Measure of Behavioral Change for Young Children with Autism. Autism: The International Journal of Research and Practice. 2003;7(1):9-30. X-1, X-3,X-4
29. Sturmey P, Lee R, Reyer H, et al. Assessing preferences for staff: Some pilot data. Behavioural and Cognitive Psychotherapy. 2003;31(1):103-7. X-1, X-3, X-4
30. Tachimori H, Osada H, Kurita H. Childhood autism rating scale--Tokyo version for screening pervasive developmental disorders. Psychiatry Clin Neurosci. 2003 Feb;57(1):113-8. PMID: 12519463; X-1
31. Webb E, Morey J, Thompsen W, et al. Prevalence of autistic spectrum disorder in children attending mainstream schools in a Welsh education authority. Dev Med Child Neurol. 2003 Jun;45(6):377-84. PMID: 12785438; X-3, X-4
32. Addington AM, Gornick M, Sporn AL, et al. Polymorphisms in the 13q33.2 gene G72/G30 are associated with childhood-onset schizophrenia and psychosis not otherwise specified. Biol Psychiatry. 2004 May 15;55(10):976-80. PMID: 15121480; X-1, X-3, X-4
33. Briggs-Gowan MJ, Carter AS, Irwin JR, et al. The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. J Pediatr Psychol. 2004 Mar;29(2):143-55. PMID: 15096535; X-1, X-3, X-4
34. Eaves LC, Ho HH. The very early identification of autism: outcome to age 4 1/2-5. J Autism Dev Disord. 2004 Aug;34(4):367-78. PMID: 15449513;
35. Geurts HM, Verte S, Oosterlaan J, et al. Can the Children's Communication Checklist differentiate between children with autism, children with ADHD, and normal controls? J Child Psychol Psychiatry. 2004 Nov;45(8):1437-53. PMID: 15482504; X-1, X-3, X-4
36. Goin PR, Myers JB. Characteristics of Infantile Autism: Moving Toward Earlier Detection. Focus on Autism and Other Developmental Disabilities. 2004;19(1):5-12. X-2, X-3, X-4, X-5
37. Howlin P, Karpf J. Using the social communication questionnaire to identify 'autistic spectrum' disorders associated with other genetic conditions: findings from a study of individuals with Cohen syndrome. Autism. 2004 Jun;8(2):175-82. PMID: 15165433; X-3, X-4
38. Munson BL. Myths and facts...about autism. Nursing. 2004 Oct;34(10):75. PMID: 15489628; X-1, X-2, X-3, X-4, X-5
39. Noland RM, Gabriels RL. Screening and identifying children with autism spectrum disorders in the public school system: the development of a model process. J Autism Dev Disord. 2004 Jun;34(3):265-77. PMID: 15264495; X-2, X-3, X-4
40. Norbury CF, Nash M, Baird G, et al. Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist--2. Int J Lang Commun Disord. 2004 Jul-Sep;39(3):345-64. PMID: 15204445; X-1, X-3,X-4
41. Rellini E, Tortolani D, Trillo S, et al. Childhood Autism Rating Scale (CARS) and Autism Behavior Checklist (ABC) correspondence and conflicts with DSM-IV criteria in diagnosis of autism. J Autism Dev Disord. 2004 Dec;34(6):703-8. PMID: 15679189; x-4
42. Skovgaard AM, Houmann T, Landorph SL, et al. Assessment and classification of psychopathology in epidemiological research of children 0-3 years of age: a review of the literature. Eur Child Adolesc Psychiatry. 2004 Dec;13(6):337-46. PMID: 15619045; X-2, X-3, X-4, X-5
43. Spence SJ, Sharifi P, Wiznitzer M. Autism spectrum disorder: screening, diagnosis, and medical evaluation. Semin Pediatr Neurol. 2004 Sep;11(3):186-95. PMID: 15575413; X-1, X-2, X-3, X-4, X-5
44. Steinhausen HC, Metzke CW. Differentiating the behavioural profile in autism and mental retardation and testing of a screener. Eur Child Adolesc Psychiatry. 2004 Aug;13(4):214-20. PMID: 15365891; X-3, X-4
45. Stone WL, Coonrod EE, Turner LM, et al. Psychometric properties of the STAT for early autism screening. J Autism Dev Disord. 2004 Dec;34(6):691-701. PMID: 15679188; X-4
46. Tebruegge M, Nandini V, Ritchie J. Does routine child health surveillance contribute to the early detection of children with pervasive developmental disorders? An epidemiological study in Kent, U.K. BMC Pediatr. 2004 Mar 3;4:4. PMID: 15053835; X-4
47. Wetherby AM, Woods J, Allen L, et al. Early indicators of autism spectrum disorders in the second year of life. J Autism Dev Disord. 2004 Oct;34(5):473-93. PMID: 15628603; X-4
48. Wong V, Hui LH, Lee WC, et al. A modified screening tool for autism (Checklist for Autism in Toddlers [CHAT-23]) for Chinese children. Pediatrics. 2004 Aug;114(2):e166-76. PMID: 15286253; X-3
49. Bennett HE, Wood CL, Hare DJ. Providing Care for adults with Autistic Spectrum Disorders in Learning Disability Services: Needs-based or Diagnosis-driven? Journal of Applied Research in Intellectual Disabilities. 2005;18(1):57-64. X-3, X-4
50. Campbell JM. Diagnostic assessment of Asperger's disorder: a review of five third-party rating scales. J Autism Dev Disord. 2005 Feb;35(1):25-35. PMID: 15796119; X-2, X-3, X-4, X-5
51. Clarke DF, Roberts W, Daraksan M, et al. The prevalence of autistic spectrum disorder in children surveyed in a tertiary care epilepsy clinic. Epilepsia. 2005 Dec;46(12):1970-7. PMID: 16393164; X-1, X-3, X-4

52. Dame Stephanie S. Challenges to autism research. J R Soc Med. 2005 Nov;98(11):523-5. PMID: 16260809; X-1, X-2, X-3, X-4, X-5
53. de Bildt A, Sytema S, Kraijer D, et al. Prevalence of pervasive developmental disorders in children and adolescents with mental retardation. Journal of Child Psychology and Psychiatry. 2005;46(3):275-86. X-1, X-3, X-4
54. Gargus RA, Yatchmink Y. Early identification and assessment of young children with autism. Med Health R I. 2005 May;88(5):147-51. PMID: 16052956; X-2, X-3, X-4, X-5
55. Glascoe FP. Screening for developmental and behavioral problems. Ment Retard Dev Disabil Res Rev. 2005;11(3):173-9. PMID: 16161092; X-2, X-3, X-4
56. Gray KM, Tonge BJ. Screening for autism in infants and preschool children with developmental delay. Australian and New Zealand Journal of Psychiatry. 2005;39(5):378-86. X-4
57. Hansson SL, Svanstrom Rojvall A, Rastam M, et al. Psychiatric telephone interview with parents for screening of childhood autism - tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): preliminary reliability and validity. Br J Psychiatry. 2005 Sep;187:262-7. PMID: 16135864; X-3, X-4
58. Honda H, Shimizu Y, Imai M, et al. Cumulative incidence of childhood autism: a total population study of better accuracy and precision. Dev Med Child Neurol. 2005 Jan;47(1):10-8. PMID: 15686284; x-4
59. Jeffries AR, Curran S, Elmslie F, et al. Molecular and phenotypic characterization of ring chromosome 22. Am J Med Genet A. 2005 Aug 30;137(2):139-47. PMID: 16059935; X-1, X-3, X-4
60. Juneja M, Mukherjee SB, Sharma S. A descriptive hospital based study of children with autism. Indian Pediatr. 2005 May;42(5):453-8. PMID: 15923691; X-4, X-5
61. Ke JY, Chen CL, Chen YJ, et al. Features of developmental functions and autistic profiles in children with fragile X syndrome. Chang Gung Med J. 2005 Aug;28(8):551-8. PMID: 16265845; X-1, X-4, X-5
62. Koegel LK, Koegel RL, Nefdt N, et al. First S.T.E.P: A Model for the Early Identification of Children With Autism Spectrum Disorders. Journal of Positive Behavior Interventions. 2005;7(4):247-52. X-2, X-3, X-4
63. Kraijer D, de Bildt A. The PDD-MRS: an instrument for identification of autism spectrum disorders in persons with mental retardation. J Autism Dev Disord. 2005 Aug;35(4):499-513. PMID: 16134035; X-4
64. Kurita H, Koyama T, Osada H. Autism-Spectrum Quotient-Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. Psychiatry Clin Neurosci. 2005 Aug;59(4):490-6. PMID: 16048456; X-3, X-4
65. Lecavalier L. An evaluation of the Gilliam Autism Rating Scale. J Autism Dev Disord. 2005 Dec;35(6):795-805. PMID: 16283084; X-3, X-4
66. Namerow LB, Mangini LM. Targeting symptom domains: a strategy for pharmacotherapy in childhood pervasive developmental disorders. Conn Med. 2005 Oct;69(9):525-33. PMID: 16270789; X-1, X-2, X-3, X-4, X-5
67. Pinto-Martin JA, Dunkle M, Earls M, et al. Developmental stages of developmental screening: steps to implementation of a successful program. Am J Public Health. 2005 Nov;95(11):1928-32. PMID: 16195523; X-2, X-3, X-4
68. Pinto-Martin JA, Souders MC, Giarelli E, et al. The role of nurses in screening for autistic spectrum disorder in pediatric primary care. J Pediatr Nurs. 2005 Jun;20(3):163-9. PMID: 15933650; X-2, X-3, X-4, X-5
69. Skuse DH, Mandy WP, Scourfield J. Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. Br J Psychiatry. 2005 Dec;187:568-72. PMID: 16319410; X-3, X-4
70. Stehouwer JD. Test Review: Rick, J.R. (2003).Krug Asperger's Disorder Index (KADI). Austin, TX: PRO-ED. Assessment for Effective Intervention. 2005;31(1):85-8. X-2, X-3, X-4, X-5
71. Towbin KE, Pradella A, Gorrindo T, et al. Autism spectrum traits in children with mood and anxiety disorders. J Child Adolesc Psychopharmacol. 2005 Jun;15(3):452-64. PMID: 16092910; X-3, X-4
72. Williams TO, Jr., Eaves RC. Pervasive developmental disorders rating scale: development and construct validity. Psychol Rep. 2005 Aug;97(1):245-57. PMID: 16279331; X-3, X-4
73. Williams TO, Eaves RC. Factor Analysis of the Pervasive Developmental Disorders Rating Scale With Teacher Ratings of Students With Autistic Disorder. Psychology in the Schools. 2005;42(2):207-16. X-4
74. Wray J, Silove N, Knott H. 7. Language disorders and autism. Med J Aust. 2005 Apr 4;182(7):354-60. PMID: 15804229; X-1, X-2, X-3, X-4, X-5
75. Zhang X, Ji CY. Autism and mental retardation of young children in China. Biomed Environ Sci. 2005 Oct;18(5):334-40. PMID: 16370317; X-1, X-4, X-5
76. Carpenter LA, Macias MM. Screening and diagnosis of autism spectrum disorders (ASD). J S C Med Assoc. 2006 Oct;102(8):271-3. PMID: 17319241; X-10
77. Crais ER, Watson LR, Baranek GT, et al. Early identification of autism: how early can we go? Semin Speech Lang. 2006 Aug;27(3):143-60. PMID: 16941286; X-2, X-3, X-4, X-5
78. Descheemaeker MJ, Govers V, Vermeulen P, et al. Pervasive developmental disorders in Prader-Willi syndrome: the Leuven experience in 59 subjects and controls. Am J Med Genet A. 2006 Jun 1;140(11):1136-42. PMID: 16646032; X-1, X-3 ,X4
79. Dosreis S, Weiner CL, Johnson L, et al. Autism spectrum disorder screening and management practices among general pediatric providers. J Dev Behav Pediatr. 2006 Apr;27(2 Suppl):S88-94. PMID: 16685190; X-3, X-4

80. Earls MF, Hay SS. Setting the stage for success: implementation of developmental and behavioral screening and surveillance in primary care practice--the North Carolina Assuring Better Child Health and Development (ABCD) Project. Pediatrics. 2006 Jul;118(1):e183-8. PMID: 16818532; X-1, X-2, X-4
81. Eaves LC, Wingert H, Ho HH. Screening for autism: agreement with diagnosis. Autism. 2006 May;10(3):229-42. PMID: 16682396;
82. Eaves LC, Wingert HD, Ho HH, et al. Screening for autism spectrum disorders with the social communication questionnaire. J Dev Behav Pediatr. 2006 Apr;27(2 Suppl):S95-S103. PMID: 16685191; X-3, X-4
83. Eaves RC, Williams TO, Jr. Exploratory and confirmatory factor analyses of the pervasive developmental disorders rating scale for young children with autistic disorder. J Genet Psychol. 2006 Mar;167(1):65-92. PMID: 16629404; X-1, X-4
84. Eaves RC, Williams TO. The reliability and construct validity of ratings for the autism behavior checklist. Psychology in the Schools. 2006;43(2):129-42. X-3,X-4
85. Eaves RC, Woods-Groves S, Williams TO, et al. Reliability and Validity of the Pervasive Developmental Disorders Rating Scale and the Gilliam Autism Rating Scale. Education and Training in Developmental Disabilities. 2006;41(3):300-9. X-3, X-4
86. Hartman CA, Luteijn E, Serra M, et al. Refinement of the Children's Social Behavior Questionnaire (CSBQ): an instrument that describes the diverse problems seen in milder forms of PDD. J Autism Dev Disord. 2006 Apr;36(3):325-42. PMID: 16617405; X-1, X-3, X-4
87. Hattori J, Ogino T, Abiru K, et al. Are pervasive developmental disorders and attention-deficit/hyperactivity disorder distinct disorders? Brain Dev. 2006 Jul;28(6):371-4. PMID: 16504439; X-1, X-3 ,X-4
88. Jane MC, Canals J, Ballespi S, et al. Parents and teachers reports of DSM-IV psychopathological symptoms in preschool children: differences between urban-rural Spanish areas. Soc Psychiatry Psychiatr Epidemiol. 2006 May;41(5):386-93. PMID: 16520884; X-1, X-3 ,X-4
89. Krol NP, De Bruyn EE, Coolen JC, et al. From CBCL to DSM: a comparison of two methods to screen for DSM-IV diagnoses using CBCL data. J Clin Child Adolesc Psychol. 2006 Feb;35(1):127-35. PMID: 16390308; X-1, X-3, X-4
90. Lord C, Risi S, DiLavore PS, et al. Autism from 2 to 9 years of age. Arch Gen Psychiatry. 2006 Jun;63(6):694-701. PMID: 16754843; X-1, X-4
91. Mawle E, Griffiths P. Screening for autism in pre-school children in primary care: systematic review of English Language tools. Int J Nurs Stud. 2006 Jul;43(5):623-36. PMID: 16387303; X-SR
92. Mazefsky CA, Oswald DP. The Discriminative Ability and Diagnostic Utility of the ADOS-G, ADI-R, and GARS for Children in a Clinical Setting. Autism: The International Journal of Research & Practice. 2006;10(6):533-49. X-4
93. Mitchell S, Brian J, Zwaigenbaum L, et al. Early language and communication development of infants later diagnosed with autism spectrum disorder. J Dev Behav Pediatr. 2006 Apr;27(2 Suppl):S69-78. PMID: 16685188; X-1, X-4
94. Persson B, Nordstrom B, Petersson K, et al. Screening for infants with developmental deficits and/or autism: a Swedish pilot study. J Pediatr Nurs. 2006 Aug;21(4):313-24. PMID: 16843217; X-4
95. Pine E, Luby J, Abbacchi A, et al. Quantitative assessment of autistic symptomatology in preschoolers. Autism. 2006 Jul;10(4):344-52. PMID: 16908478; X-1, X-3, X-4
96. Risi S, Lord C, Gotham K, et al. Combining Information from Multiple Sources in the Diagnosis of Autism Spectrum Disorders. Journal of the American Academy of Child and Adolescent Psychiatry. 2006;45(9):1094-103. X-1, X-3,X-4
97. Robins DL, Dumont-Mathieu TM. Early screening for autism spectrum disorders: update on the modified checklist for autism in toddlers and other measures. J Dev Behav Pediatr. 2006 Apr;27(2 Suppl):S111-9. PMID: 16685177; X-2, X-3, X-4, X-5
98. Scambler DJ, Hepburn SL, Rogers SJ. A two-year follow-up on risk status identified by the checklist for autism in toddlers. J Dev Behav Pediatr. 2006 Apr;27(2 Suppl):S104-10. PMID: 16685176; X-4
99. Swinkels SH, Dietz C, van Daalen E, et al. Screening for autistic spectrum in children aged 14 to 15 months. I: the development of the Early Screening of Autistic Traits Questionnaire (ESAT). J Autism Dev Disord. 2006 Aug;36(6):723-32. PMID: 16614790; X-4
100. Tryon PA, Mayes SD, Rhodes RL, et al. Can Asperger's Disorder Be Differentiated from Autism Using "DSM-IV" Criteria? Focus on Autism and Other Developmental Disabilities. 2006;21(1):2-6. X-1, X-2, X-3, X-4
101. Williams J, Brayne C. Screening for autism spectrum disorders. What is the evidence? Autism. 2006;10(1):11-35. X-2, X-3, X-4
102. Allen CW, Silove N, Williams K, et al. Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. J Autism Dev Disord. 2007 Aug;37(7):1272-8. PMID: 17080270;
103. Allison C, Williams J, Scott F, et al. The Childhood Asperger Syndrome Test (CAST): test-retest reliability in a high scoring sample. Autism. 2007 Mar;11(2):173-85. PMID: 17353217; X-3, X-4
104. Almon-Morris H, Diakite A. Teaching Emotions to Children with Autism: Identification, Demonstration, and Explanation of Occasioning Stimuli. Journal of Precision Teaching and Celeration. 2007;23:19-22. X-1, X-3, X-4
105. Angley M, Young R, Ellis D, et al. Children and autism--Part 1--recognition and pharmacological management. Aust Fam Physician. 2007 Sep;36(9):741-4. PMID: 17915375; X-1, X-2, X-3, X-4
106. Ben-Sasson A, Cermak SA, Orsmond GI, et al. Extreme sensory modulation behaviors in toddlers with autism spectrum disorders. Am J Occup Ther. 2007 Sep-Oct;61(5):584-92. PMID: 17944296; X-1, X-4

Appendix E. Excluded Studies

107. Bruckner C, Yoder P, Stone W, et al. Construct validity of the MCDI-I Receptive Vocabulary scale can be improved: differential item functioning between toddlers with autism spectrum disorders and typically developing infants. J Speech Lang Hear Res. 2007 Dec;50(6):1631-8. PMID: 18055777; X-1, X-4
108. Bryson SE, Zwaigenbaum L, Brian J, et al. A prospective case series of high-risk infants who developed autism. J Autism Dev Disord. 2007 Jan;37(1):12-24. PMID: 17211728; X-1, X-3, X-4
109. Chawarska K, Klin A, Paul R, et al. Autism spectrum disorder in the second year: stability and change in syndrome expression. J Child Psychol Psychiatry. 2007 Feb;48(2):128-38. PMID: 17300551; X-1, X-4
110. Clifford S, Young R, Williamson P. Assessing the early characteristics of autistic disorder using video analysis. J Autism Dev Disord. 2007 Feb;37(2):301-13. PMID: 17031450; X-1, X-2, X-4
111. Corsello C, Hus V, Pickles A, et al. Between a ROC and a hard place: decision making and making decisions about using the SCQ. J Child Psychol Psychiatry. 2007 Sep;48(9):932-40. PMID: 17714378;
112. Eapen V, Mabrouk AA, Zoubeidi T, et al. Prevalence of pervasive developmental disorders in preschool children in the UAE. J Trop Pediatr. 2007 Jun;53(3):202-5. PMID: 17244665; X-3
113. Ellefsen A, Kampmann H, Billstedt E, et al. Autism in the Faroe Islands: an epidemiological study. J Autism Dev Disord. 2007 Mar;37(3):437-44. PMID: 17029020; X-1, X-3, X-4, X-5
114. Glascoe FP, Macias MM, Wegner LM, et al. Can a broadband developmental-behavioral screening test identify children likely to have autism spectrum disorder? Clin Pediatr (Phila). 2007 Nov;46(9):801-5. PMID: 17641122; X-4

RETAIN FOR CQ

115. Gotham K, Risi S, Pickles A, et al. The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. J Autism Dev Disord. 2007 Apr;37(4):613-27. PMID: 17180459; X-1, X-4
116. Groen WB, Swinkels SH, van der Gaag RJ, et al. Finding effective screening instruments for autism using bayes theorem. Arch Pediatr Adolesc Med. 2007 Apr;161(4):415-6. PMID: 17404143; X-4
117. Lee L-C, David AB, Rusyniak J, et al. Performance of the Social Communication Questionnaire in children receiving preschool special education services. Research in Autism Spectrum Disorders. 2007;1(2):126-38. X-4
118. Limperopoulos C, Bassan H, Gauvreau K, et al. Does cerebellar injury in premature infants contribute to the high prevalence of long-term cognitive, learning, and behavioral disability in survivors? Pediatrics. 2007 Sep;120(3):584-93. PMID: 17766532; X-4
119. Magyar CI, Pandolfi V. Factor structure evaluation of the childhood autism rating scale. J Autism Dev Disord. 2007 Oct;37(9):1787-94. PMID: 17437070; X-1, X-4
120. Mattila ML, Kielinen M, Jussila K, et al. An epidemiological and diagnostic study of Asperger syndrome according to four sets of diagnostic criteria. J Am Acad Child Adolesc Psychiatry. 2007 May;46(5):636-46. PMID: 17450055; X-1, X-2, X-3, X-4
121. McClure I, Melville CA. Early identification key in autism spectrum disorders. Practitioner. 2007 Aug;251(1697):31, 3-5. PMID: 17882895; X-1, X-2, X-3, X-4
122. Nadel S, Poss JE. Early detection of autism spectrum disorders: screening between 12 and 24 months of age. J Am Acad Nurse Pract. 2007 Aug;19(8):408-17. PMID: 17655570; X-1, X-2, X-3, X-4
123. Nadig AS, Ozonoff S, Young GS, et al. A prospective study of response to name in infants at risk for autism. Arch Pediatr Adolesc Med. 2007 Apr;161(4):378-83. PMID: 17404135;
124. Reznick JS, Baranek GT, Reavis S, et al. A parent-report instrument for identifying one-year-olds at risk for an eventual diagnosis of autism: The First Year Inventory. Journal of Autism and Developmental Disorders. 2007;37(9):1691-710. X-4
125. Richler J, Bishop SL, Kleinke JR, et al. Restricted and repetitive behaviors in young children with autism spectrum disorders. J Autism Dev Disord. 2007 Jan;37(1):73-85. PMID: 17195920; X-1, X-4
126. Saemundsen E, Ludvigsson P, Hilmarsdottir I, et al. Autism spectrum disorders in children with seizures in the first year of life - a population-based study. Epilepsia. 2007 Sep;48(9):1724-30. PMID: 17555525; X-3, X-4
127. Scambler DJ, Hepburn S, Rutherford MD, et al. Emotional responsivity in children with autism, children with other developmental disabilities, and children with typical development. J Autism Dev Disord. 2007 Mar;37(3):553-63. PMID: 16933089; X-1, X-4
128. Scambler DJ, Hepburn SL, Hagerman RJ, et al. A preliminary study of screening for risk of autism in children with fragile X syndrome: testing two risk cut-offs for the Checklist for Autism in Toddlers. J Intellect Disabil Res. 2007 Apr;51(Pt 4):269-76. PMID: 17326808; X-4
129. Stone WL, McMahon CR, Yoder PJ, et al. Early social-communicative and cognitive development of younger siblings of children with autism spectrum disorders. Arch Pediatr Adolesc Med. 2007 Apr;161(4):384-90. PMID: 17404136; X-4
130. Ventola P, Kleinman J, Pandey J, et al. Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. J Autism Dev Disord. 2007 Mar;37(3):425-36. PMID: 16897377; X-4
131. Watson LR, Baranek GT, Crais ER, et al. The first year inventory: retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. J Autism Dev Disord. 2007 Jan;37(1):49-61. PMID: 17219058; X-4
132. Wiggins LD, Bakeman R, Adamson LB, et al. The utility of the Social Communication Questionnaire in screening for autism in children referred for early intervention. Focus on Autism and Other Developmental Disabilities. 2007;22(1):33-8. X-4
133. Witwer AN, Lecavalier L. Autism screening tools: an evaluation of the Social Communication Questionnaire and the Developmental Behaviour Checklist-Autism Screening Algorithm. J Intellect Dev Disabil. 2007 Sep;32(3):179-87. PMID: 17885895; X-3, X-4

134. Allen RA, Robins DL, Decker SL. Autism spectrum disorders: Neurobiology and current assessment practices. Psychology in the Schools. 2008;45(10):905-17. X-2, X-3, X-4, X-5
135. Allison C, Baron-Cohen S, Wheelwright S, et al. The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): a normally distributed quantitative measure of autistic traits at 18-24 months of age: preliminary report. J Autism Dev Disord. 2008 Sep;38(8):1414-25. PMID: 18240013; X-4
136. Bolte S, Holtmann M, Poustka F. The Social Communication Questionnaire (SCQ) as a screener for autism spectrum disorders: additional evidence and cross-cultural validity. J Am Acad Child Adolesc Psychiatry. 2008 Jun;47(6):719-20; author reply 20-1. PMID: 18496329; X-3, X-4
137. Boomsma A, Van Lang ND, De Jonge MV, et al. A new symptom model for autism cross-validated in an independent sample. J Child Psychol Psychiatry. 2008 Aug;49(8):809-16. PMID: 18492042; X-1, X-3, X-4
138. Brian J, Bryson SE, Garon N, et al. Clinical assessment of autism in high-risk 18-month-olds. Autism. 2008 Sep;12(5):433-56. PMID: 18805941; X-1, X-3, X-4
139. Bryson SE, Zwaigenbaum L, McDermott C, et al. The Autism Observation Scale for Infants: scale development and reliability data. J Autism Dev Disord. 2008 Apr;38(4):731-8. PMID: 17874180; X-4
140. Chiang CH, Soong WT, Lin TL, et al. Nonverbal communication skills in young children with autism. J Autism Dev Disord. 2008 Nov;38(10):1898-906. PMID: 18491223; X-1, X-4,X-5
141. DeVincent C, Gadow KD, Strong G, et al. Screening for autism spectrum disorder with the Early Childhood Inventory-4. J Dev Behav Pediatr. 2008 Feb;29(1):1-10. PMID: 18300725; X-3, X-4
142. Elder LM, Dawson G, Toth K, et al. Head circumference as an early predictor of autism symptoms in younger siblings of children with autism spectrum disorder. J Autism Dev Disord. 2008 Jul;38(6):1104-11. PMID: 18058011; X-4
143. Farmer JE, Clark MJ. Identification and evaluation of Missouri's children with autism spectrum disorders: promoting a rapid response. Mo Med. 2008 Sep-Oct;105(5):384-9. PMID: 18807359; X-2, X-3, X-4, X-5
144. Garcia-Primo P, Martin-Arribas MC, Ferrari-Arroyo MJ, et al. Autism, the big unkown. Georgian Med News. 2008 Mar(156):9-14. PMID: 18403803; X-2, X-3,X-4
145. Ghanizadeh A. A preliminary study on screening prevalence of pervasive developmental disorder in schoolchildren in Iran. J Autism Dev Disord. 2008 Apr;38(4):759-63. PMID: 17879153; X-3, X-4, X-5
146. Goin-Kochel RP, Cohen R. Screening cases within a statewide autism registry: A comparison of parental reports using DSM-IV-TR criteria versus the SCQ. Focus on Autism and Other Developmental Disabilities. 2008;23(3):148-54. X-3,X-4
147. Gray KM, Tonge BJ, Sweeney DJ, et al. Screening for autism in young children with developmental delay: an evaluation of the developmental behaviour checklist: early screen. J Autism Dev Disord. 2008 Jul;38(6):1003-10. PMID: 17972169;
148. Greenspan SI, Brazelton TB, Cordero J, et al. Guidelines for early identification, screening, and clinical management of children with autism spectrum disorders. Pediatrics. 2008 Apr;121(4):828-30. PMID: 18381546; X-2, X-3, X-4, X-5
149. Herba CM, de Bruin E, Althaus M, et al. Face and emotion recognition in MCDD versus PDD-NOS. J Autism Dev Disord. 2008 Apr;38(4):706-18. PMID: 17721812; X-1, X-3, X-4
150. Hess KL, Morrier MJ, Heflin LJ, et al. Autism treatment survey: services received by children with autism spectrum disorders in public school classrooms. J Autism Dev Disord. 2008 May;38(5):961-71. PMID: 17929155; X-1, X-3, X-4
151. Hoekstra RA, Bartels M, Cath DC, et al. Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. J Autism Dev Disord. 2008 Sep;38(8):1555-66. PMID: 18302013; X-1, X-3,X-4
152. Johnson CP. Recognition of autism before age 2 years. Pediatr Rev. 2008 Mar;29(3):86-96. PMID: 18310467; X-2, X-3, X-4, X-5
153. Kawamura Y, Takahashi O, Ishii T. Reevaluating the incidence of pervasive developmental disorders: Impact of elevated rates of detection through implementation of an integrated system of screening in Toyota, Japan. Psychiatry and Clinical Neurosciences. 2008;62(2):152-9. X-4
154. Kleinman JM, Ventola PE, Pandey J, et al. Diagnostic stability in very young children with autism spectrum disorders. J Autism Dev Disord. 2008 Apr;38(4):606-15. PMID: 17924183; X-4
155. Kurita H, Koyama T, Inoue K. Reliability and validity of the Pervasive Developmental Disorders Assessment System. Psychiatry Clin Neurosci. 2008 Apr;62(2):226-33. PMID: 18412847; X-1, X-4
156. Lejarraga H, Menendez AM, Menzano E, et al. Screening for developmental problems at primary care level: a field programme in San Isidro, Argentina. Paediatr Perinat Epidemiol. 2008 Mar;22(2):180-7. PMID: 18298693; X-1, X-3
157. Limperopoulos C, Bassan H, Sullivan NR, et al. Positive screening for autism in ex-preterm infants: prevalence and risk factors. Pediatrics. 2008 Apr;121(4):758-65. PMID: 18381541; X-4

RETAIN FOR CONTEXT

158. Luyster RJ, Kadlec MB, Carter A, et al. Language assessment and development in toddlers with autism spectrum disorders. J Autism Dev Disord. 2008 Sep;38(8):1426-38. PMID: 18188685; X-1, X-4
159. Matson JL, Dempsey T, Rivet T. A comparison of Asperger symptom rating scales with children and adolescents. Research in Autism Spectrum Disorders. 2008;2(4):643-50. X-3
160. Matson JL, Gonzalez ML, Rivet TT. Reliability of the Autism Spectrum Disorder-Behavior Problems for Children (ASD-BPC). Research in Autism Spectrum Disorders. 2008;2(4):696-706. X-1, X-4
161. Ozonoff S, Heung K, Byrd R, et al. The onset of autism: patterns of symptom emergence in the first years of life. Autism Res. 2008 Dec;1(6):320-8. PMID: 19360687; X-1, X-2, X-3, X-4

Appendix E. Excluded Studies

162. Ozonoff S, Macari S, Young GS, et al. Atypical object exploration at 12 months of age is associated with autism in a prospective sample. Autism. 2008 Sep;12(5):457-72. PMID: 18805942; X-1, X-3, X-4
163. Pine DS, Guyer AE, Goldwin M, et al. Autism spectrum disorder scale scores in pediatric mood and anxiety disorders. J Am Acad Child Adolesc Psychiatry. 2008 Jun;47(6):652-61. PMID: 18434923; X-1, X-3,X-4
164. Pinto-Martin JA, Young LM, Mandell DS, et al. Screening strategies for autism spectrum disorders in pediatric primary care. J Dev Behav Pediatr. 2008 Oct;29(5):345-50. PMID: 18852608; X-4
165. Rapin I, Tuchman RF. Autism: definition, neurobiology, screening, diagnosis. Pediatr Clin North Am. 2008 Oct;55(5):1129-46, viii. PMID: 18929056; X-2, X-3, X-4, X-5
166. Ronald A, Happe F, Plomin R. A twin study investigating the genetic and environmental aetiologies of parent, teacher and child ratings of autistic-like traits and their overlap. Eur Child Adolesc Psychiatry. 2008 Dec;17(8):473-83. PMID: 18427861; X-1, X-3,X-4
167. Santaella ML, Varela Y, Linares N, et al. Prevalence of autism spectrum disorders in relatives of patients with selective immunoglobulin A deficiency. P R Health Sci J. 2008 Sep;27(3):204-8. PMID: 18782963; X-1, X-3, X-4, X-5
168. Scholte EM, Van Berckelaer-Onnes I, Van der Ploeg JD. Rating scale to screen symptoms of psychiatric disorders in children. European Journal of Special Needs Education. 2008;23(1):47-62. X-3,X-4
169. Seif Eldin A, Habib D, Noufal A, et al. Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. Int Rev Psychiatry. 2008 Jun;20(3):281-9. PMID: 18569180; X-5
170. Sikora DM, Hall TA, Hartley SL, et al. Does parent report of behavior differ across ADOS-G classifications: analysis of scores from the CBCL and GARS. J Autism Dev Disord. 2008 Mar;38(3):440-8. PMID: 17619131; X-3, X-4
171. Snow AV, Lecavalier L. Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. Autism. 2008 Nov;12(6):627-44. PMID: 19005032;
172. Steyaert JG, De la Marche W. What's new in autism? Eur J Pediatr. 2008 Oct;167(10):1091-101. PMID: 18597114; X-1, X-2, X-3, X-4, X-5
173. Stone WL, McMahon CR, Henderson LM. Use of the Screening Tool for Autism in Two-Year-Olds (STAT) for children under 24 months: an exploratory study. Autism. 2008 Sep;12(5):557-73. PMID: 18805947; X-4
174. Suhail K, Zafar F. Prevalence of autism in special education schools of Lahore. Pakistan Journal of Psychological Research. 2008;23(3-4):45-64. X-1, X-4, X-5
175. Taniai H, Nishiyama T, Miyachi T, et al. Genetic influences on the broad spectrum of autism: study of proband-ascertained twins. Am J Med Genet B Neuropsychiatr Genet. 2008 Sep 5;147B(6):844-9. PMID: 18361421; X-1, X-3, X-4
176. Wallis KE, Pinto-Martin J. The challenge of screening for autism spectrum disorder in a culturally diverse society. Acta Paediatr. 2008 May;97(5):539-40. PMID: 18373717; X-2, X-3,X-4
177. Wallis KE, Smith SM. School health developmental screening in pediatric primary care: the role of nurses. J Spec Pediatr Nurs. 2008 Apr;13(2):130-4. PMID: 18366381; X-2, X-3, X-4, X-5
178. Wiggins LD, Robins DL. Brief report: excluding the ADI-R behavioral domain improves diagnostic agreement in toddlers. J Autism Dev Disord. 2008 May;38(5):972-6. PMID: 17879150; X-4
179. Anderson GM. Conceptualizing Autism: The Role for Emergence. Journal of the American Academy of Child & Adolescent Psychiatry. 2009;48(7):688-. X-1, X-2, X-3, X-4
180. Bauer SC, Smith PJ, Chien AT, et al. Educating Pediatric Residents about Developmental and Social-Emotional Health. Infants and Young Children. 2009;22(4):309-20. X-1, X-3, X-4, X-5
181. Bhattacharyya R, Sanyal D, Roy K, et al. A study of cluster behavioral abnormalities in Down syndrome. Indian J Med Sci. 2009 Feb;63(2):58-65. PMID: 19359768; X-1, X-3, X-4
182. Chakrabarti S. Early identification of autism. Indian Pediatr. 2009 May;46(5):412-4. PMID: 19179745; X-1, X-3, X-4, X-5
183. de Bildt A, Mulder EJ, Hoekstra PJ, et al. Validity of the Children's Social Behavior Questionnaire (CSBQ) in children with intellectual disability: comparing the CSBQ with ADI-R, ADOS, and clinical DSM-IV-TR classification. J Autism Dev Disord. 2009 Oct;39(10):1464-70. PMID: 19495951; X-1, X-2, X-3, X-4
184. Deb S, Dhaliwal A-J, Roy M. The usefulness of the DBC-ASA as a screening instrument for autism in children with intellectual disabilities: A pilot study. Journal of Applied Research in Intellectual Disabilities. 2009;22(5):498-501. X-3
185. Duby JC, Johnson CP. Universal screening for autism spectrum disorders: a snapshot within the big picture. Pediatr Ann. 2009 Jan;38(1):36-41. PMID: 19213292; X-2, X-3 ,X-4
186. Earls MF, Andrews JE, Hay SS. A longitudinal study of developmental and behavioral screening and referral in North Carolina's Assuring Better Child Health and Development participating practices. Clin Pediatr (Phila). 2009 Oct;48(8):824-33. PMID: 19571336; X-1
187. Fodstad JC, Matson JL, Hess J, et al. Social and communication behaviours in infants and toddlers with autism and pervasive developmental disorder-not otherwise specified. Dev Neurorehabil. 2009 Jun;12(3):152-7. PMID: 19466623; X-4
188. Gillis JM. Screening practices of family physicians and pediatricians in 2 southern states. Infants & Young Children. 2009;22(4):321-31. X-3, X-4
189. Goldson E. Autism: an update. Adv Pediatr. 2009;56:187-201. PMID: 19968949; X-2, X-3, X-4, X-5
190. Goldwater- Adler S, MacLeod CG, Smith IM. DSMD in triage of youth seen at a specialty autism clinic. Journal on Developmental Disabilities. 2009;15(2):63-71. X-1, X-3, X-4

Appendix E. Excluded Studies

191. Hack M, Taylor HG, Schluchter M, et al. Behavioral outcomes of extremely low birth weight children at age 8 years. J Dev Behav Pediatr. 2009 Apr;30(2):122-30. PMID: 19322106; X-1, X-3 ,X-4
192. Hepburn SL, Maclean WE. Maladaptive and repetitive behaviors in children with Down syndrome and autism spectrum disorders: Implications for screening. Journal of Mental Health Research in Intellectual Disabilities. 2009;2(2):67-88. X-1, X-3 ,X-4
193. Inglese MD. Caring for children with autism spectrum disorder. Part II: screening, diagnosis, and management. J Pediatr Nurs. 2009 Feb;24(1):49-59. PMID: 19159835; X-2, X-3, X-4, X-5
194. Karabekiroglu K, Aman MG. Validity of the aberrant behavior checklist in a clinical sample of toddlers. Child Psychiatry Hum Dev. 2009 Mar;40(1):99-110. PMID: 18600444; X-1, X-4 ,X-5
195. Kiser K. The power of books. Minn Med. 2009 Mar;92(3):9-11. PMID: 19400377; X-1, X-2, X-3, X-4
196. Kuban KC, O'Shea TM, Allred EN, et al. Positive screening on the Modified Checklist for Autism in Toddlers (M-CHAT) in extremely low gestational age newborns. J Pediatr. 2009 Apr;154(4):535-40 e1. PMID: 19185317; X-4
197. Levy SE, Mandell DS, Schultz RT. Autism. Lancet. 2009 Nov 7;374(9701):1627-38. PMID: 19819542; X-2, X-3, X-4, X-5
198. Limperopoulos C. Autism spectrum disorders in survivors of extreme prematurity. Clin Perinatol. 2009 Dec;36(4):791-805, vi. PMID: 19944836; X-2, X-3, X-4, X-5
199. Matson JL, Boisjoli J, Rojahn J, et al. A factor analysis of challenging behaviors assessed with the baby and infant screen for children with autism traits (Biscuit-Part 3). Research in Autism Spectrum Disorders. 2009;3(3):714-22. X-1, X-4
200. Matson JL, Dempsey T, Fodstad JC. Stereotypies and repetitive/restrictive behaviours in infants with autism and pervasive developmental disorder. Dev Neurorehabil. 2009 Jun;12(3):122-7. PMID: 19466619; X-4
201. Matson JL, Fodstad JC, Dempsey T. What symptoms predict the diagnosis of autism or PDD-NOS in infants and toddlers with developmental delays using the Baby and Infant Screen for aUtIsm Traits. Dev Neurorehabil. 2009;12(6):381-8. PMID: 20205546;
202. Matson JL, Fodstad JC, Mahan S. Cutoffs, norms, and patterns of comorbid difficulties in children with developmental disabilities on the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT-Part 2). Res Dev Disabil. 2009 Nov-Dec;30(6):1221-8. PMID: 19450951; X-1, X-3, X-4
203. Matson JL, Wilkins J, Sevin JA, et al. Reliability and item content of the baby and infant screen for children with aUtIsm traits (BISCUIT): Parts 1–3. Research in Autism Spectrum Disorders. 2009;3(2):336-44. X-4
204. Matson JL, Wilkins J, Sharp B, et al. Sensitivity and specificity of the Baby and Infant Screen for Children with Autism Traits (BISCUIT): Validity and cutoff scores for autism and PDD-NOS in toddlers. Research in Autism Spectrum Disorders. 2009;3(4):924-30.
205. Niklasson L, Rasmussen P, Óskarsdóttir S, et al. Autism, ADHD, mental retardation and behavior problems in 100 individuals with 22q11 deletion syndrome. Research in Developmental Disabilities. 2009;30(4):763-73. X-1
206. Oosterling IJ, Swinkels SH, van der Gaag RJ, et al. Comparative analysis of three screening instruments for autism spectrum disorder in toddlers at high risk. J Autism Dev Disord. 2009 Jun;39(6):897-909. PMID: 19205862;
207. Pandolfi V, Magyar CI, Dill CA. Confirmatory factor analysis of the child behavior checklist 1.5-5 in a sample of children with autism spectrum disorders. J Autism Dev Disord. 2009 Jul;39(7):986-95. PMID: 19263208; X-1, X-4
208. Perera H, Wijewardena K, Aluthwelage R. Screening of 18-24-month-old children for autism in a semi-urban community in Sri Lanka. J Trop Pediatr. 2009 Dec;55(6):402-5. PMID: 19401407; X-5
209. Rojahn J, Matson JL, Mahan S, et al. Cutoffs, norms, and patterns of problem behaviors in children with an ASD on the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT-Part 3). Research in Autism Spectrum Disorders. 2009;3(4):989-98. X-1, X-4
210. Roscoe EM, Rooker GW, Pence ST, et al. Assessing the utility of a demand assessment for functional analysis. J Appl Behav Anal. 2009 Winter;42(4):819-25. PMID: 20514188; X-1, X-3, X-4
211. Santosh PJ, Mandy WP, Puura K, et al. The construction and validation of a short form of the developmental, diagnostic and dimensional interview. Eur Child Adolesc Psychiatry. 2009 Aug;18(8):521-4. PMID: 19263189; X-1, X-3, X-4
212. Sato FP, Paula CS, Lowenthal R, et al. Instrument to screen cases of pervasive developmental disorder: a preliminary indication of validity. Rev Bras Psiquiatr. 2009 Mar;31(1):30-3. PMID: 19506772; X-3, X-4, X-5
213. Shattuck PT, Durkin M, Maenner M, et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. J Am Acad Child Adolesc Psychiatry. 2009 May;48(5):474-83. PMID: 19318992; X-1, X-3, X-4
214. van Daalen E, Kemner C, Dietz C, et al. Inter-rater reliability and stability of diagnoses of autism spectrum disorder in children identified through screening at a very young age. Eur Child Adolesc Psychiatry. 2009 Nov;18(11):663-74. PMID: 19421728; X-4
215. Van Dyke EM. Autistic disorder: early interventions can improve outcomes. JAAPA. 2009 Jul;22(7):18-9. PMID: 19697566; X-1, X-2, X-3, X-4
216. Warren Z, Stone W, Humberd Q. A training model for the diagnosis of autism in community pediatric practice. J Dev Behav Pediatr. 2009 Oct;30(5):442-6. PMID: 19823138; X-4
217. Wiggins LD, Robins DL, Bakeman R, et al. Brief report: sensory abnormalities as distinguishing symptoms of autism spectrum disorders in young children. J Autism Dev Disord. 2009 Jul;39(7):1087-91. PMID: 19283461; X-1, X-4

218. Xu D, Gilkerson J, Richards J, et al. Child vocalization composition as discriminant information for automatic autism detection. Conf Proc IEEE Eng Med Biol Soc. 2009;2009:2518-22. PMID: 19964971; X-1, X-3, X-4
219. Zwaigenbaum L, Bryson S, Lord C, et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. Pediatrics. 2009 May;123(5):1383-91. PMID: 19403506; X-2, X-3, X-4, X-5
220. Allen SG, Berry AD, Brewster JA, et al. Enhancing developmentally oriented primary care: an Illinois initiative to increase developmental screening in medical homes. Pediatrics. 2010 Dec;126 Suppl 3:S160-4. PMID: 21123480; X-1,X-3,X-4
221. Biederman J, Petty CR, Fried R, et al. Child behavior checklist clinical scales discriminate referred youth with autism spectrum disorder: a preliminary study. J Dev Behav Pediatr. 2010 Jul-Aug;31(6):485-90. PMID: 20585266; X-3, X-4
222. Charman T. Developmental approaches to understanding and treating autism. Folia Phoniatr Logop. 2010;62(4):166-77. PMID: 20460929; X-1, X-2, X-3, X-4
223. Chlebowski C, Green JA, Barton ML, et al. Using the Childhood Autism Rating Scale to Diagnose Autism Spectrum Disorders. Journal of Autism and Developmental Disorders. 2010;40(7):787-99. X-4
224. Dereu M, Warreyn P, Raymaekers R, et al. Screening for autism spectrum disorders in Flemish day-care centres with the checklist for early signs of developmental disorders. J Autism Dev Disord. 2010 Oct;40(10):1247-58. PMID: 20198413; X-1
225. DiGuiseppi C, Hepburn S, Davis JM, et al. Screening for autism spectrum disorders in children with Down syndrome: population prevalence and screening test characteristics. J Dev Behav Pediatr. 2010 Apr;31(3):181-91. PMID: 20375732; X-4
226. Fujibayashi H, Kitayama S, Matsuo M. Score of inattention subscale of ADHD rating scale-IV is significantly higher for AD/HD than PDD. Kobe J Med Sci. 2010;56(1):E12-7. PMID: 21063141; X-3, X-4
227. Ghaziuddin M, Welch K, Mohiuddin S, et al. Utility of the Social and Communication Questionnaire in the differentiation of autism from ADHD. Journal of Developmental and Physical Disabilities. 2010;22(4):359-66. X-3, X-4, X-5
228. Gillberg C. The ESSENCE in child psychiatry: Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations. Res Dev Disabil. 2010 Nov-Dec;31(6):1543-51. PMID: 20634041; X-1, X-2, X-3, X-4
229. Granader YE, Bender HA, Zemon V, et al. The clinical utility of the Social Responsiveness Scale and Social Communication Questionnaire in tuberous sclerosis complex. Epilepsy Behav. 2010 Jul;18(3):262-6. PMID: 20554253; X-3, X-4
230. Gunnar MR. A Commentary on "Deprivation-Specific Psychological Patterns: Effects of Institutional Deprivation". Monographs of the Society for Research in Child Development. 2010;75(1):232-47. X-1, X-2, X-3, X-4
231. Hedley D, Young R, Angelica M, et al. Cross-cultural evaluation of the Autism Detection in Early Childhood (ADEC) in Mexico. Autism. 2010 Mar;14(2):93-112. PMID: 20395280; X-5
232. Hoglund Carlsson L, Gillberg C, Lannero E, et al. Autism: screening toddlers with CHAT in a child health care programme did not improve early identification. Acta Paediatr. 2010 Dec;99(12):1897-9. PMID: 20670307; X-4
233. Horovitz M, Matson JL. Communication deficits in babies and infants with autism and pervasive developmental disorder-not otherwise specified (PDD-NOS). Dev Neurorehabil. 2010;13(6):390-8. PMID: 20887201; X-1, X-4
234. Hutman T, Rozga A, DeLaurentis AD, et al. Response to distress in infants at risk for autism: a prospective longitudinal study. J Child Psychol Psychiatry. 2010 Sep;51(9):1010-20. PMID: 20546081; X-1
235. Inada N, Kamio Y, Koyama T. Developmental chronology of preverbal social behaviors in infancy using the M-CHAT: Baseline for early detection of atypical social development. Research in Autism Spectrum Disorders. 2010;4(4):605-11. X-3, X-4
236. Juneja M, Sharma S, Mukherjee SB. Sensitivity of the Autism Behavior Checklist in Indian autistic children. Journal of Developmental and Behavioral Pediatrics. 2010;31(1):48-9. X-1, X-3, X-4, X-5
237. Karabekiroglu K, Briggs-Gowan MJ, Carter AS, et al. The clinical validity and reliability of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). Infant Behav Dev. 2010 Dec;33(4):503-9. PMID: 20800285; X-5
238. Karmel BZ, Gardner JM, Meade LS, et al. Early medical and behavioral characteristics of NICU infants later classified with ASD. Pediatrics. 2010 Sep;126(3):457-67. PMID: 20679296; X-1, X-3, X-4
239. Koyama T, Inokuchi E, Inada N, et al. Utility of the Japanese version of the Checklist for Autism in Toddlers for predicting pervasive developmental disorders at age 2. Psychiatry Clin Neurosci. 2010 Jun;64(3):330-2. PMID: 20602732;
240. Larson T, Anckarsater H, Gillberg C, et al. The autism--tics, AD/HD and other comorbidities inventory (A-TAC): further validation of a telephone interview for epidemiological research. BMC Psychiatry. 2010;10:1. PMID: 20055988; X-3, X-4
241. Lowenthal R, Mercadante MT, Filho JFB, et al. Autism spectrum disorder in Down syndrome: Definition of the cutoff point for the Autism Screening Questionnaire screening instrument. Journal of Developmental and Behavioral Pediatrics. 2010;31(8):684-. X-2, X-3, X-4, X-5
242. Lung FW, Shu BC, Chiang TL, et al. Parental concerns based general developmental screening tool and autism risk: the Taiwan National Birth cohort study. Pediatr Res. 2010 Feb;67(2):226-31. PMID: 20091940; X-1, X-3, X-4, X-5
243. Matson JL, Boisjoli JA, Hess JA, et al. Factor structure and diagnostic fidelity of the Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT-part 1). Dev Neurorehabil. 2010;13(2):72-9. PMID: 20222767; X-4

Appendix E. Excluded Studies

244. Matson JL, Fodstad JC, Mahan S, et al. Cut-offs, norms and patterns of problem behaviours in children with developmental disabilities on the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT-Part 3). Dev Neurorehabil. 2010 Feb;13(1):3-9. PMID: 20067340; X-1, X-3, X-4
245. McClure I, MacKay T, Mamdani H, et al. A Comparison of a Specialist Autism Spectrum Disorder Assessment Team with Local Assessment Teams. Autism: The International Journal of Research and Practice. 2010;14(6):589-603. X-1, X-3, X-4
246. Moricke E, Swinkels SH, Beuker KT, et al. Predictive value of subclinical autistic traits at age 14-15 months for behavioural and cognitive problems at age 3-5 years. Eur Child Adolesc Psychiatry. 2010 Aug;19(8):659-68. PMID: 20390313; X-4
247. Oosterling I, Rommelse N, de Jonge M, et al. How useful is the Social Communication Questionnaire in toddlers at risk of autism spectrum disorder? J Child Psychol Psychiatry. 2010 Nov;51(11):1260-8. PMID: 20626528;
248. Oosterling IJ, Wensing M, Swinkels SH, et al. Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. J Child Psychol Psychiatry. 2010 Mar;51(3):250-8. PMID: 19843319;
249. Ozonoff S, Iosif AM, Baguio F, et al. A prospective study of the emergence of early behavioral signs of autism. J Am Acad Child Adolesc Psychiatry. 2010 Mar;49(3):256-66 e1-2. PMID: 20410715; X-1, X-3, X-4
250. Pandolfi V, Magyar CI, Dill CA. Constructs assessed by the GARS-2: factor analysis of data from the standardization sample. J Autism Dev Disord. 2010 Sep;40(9):1118-30. PMID: 20151186; X-3, X-4
251. Rao S, Salmon G. Autism spectrum disorders. Br J Hosp Med (Lond). 2010 Dec;71(12):699-703. PMID: 21135768; X-1, X-2, X-3, X-4, X-5
252. Rodman JL, Gilbert KA, Grove AB, et al. Efficacy of brief quantitative measures of play for screening for autism spectrum disorders. J Autism Dev Disord. 2010 Mar;40(3):325-33. PMID: 19798562; X-1, X-4
253. Staines R. School nurses can help identify children with undiagnosed autism. Paediatr Nurs. 2010 Mar;22(2):7. PMID: 20373658; X-1, X-2, X-3, X-4
254. Stilp RL, Gernsbacher MA, Schweigert EK, et al. Genetic variance for autism screening items in an unselected sample of toddler-age twins. J Am Acad Child Adolesc Psychiatry. 2010 Mar;49(3):267-76. PMID: 20410716; X-4
255. Stoltenberg C, Schjolberg S, Bresnahan M, et al. The Autism Birth Cohort: a paradigm for gene-environment-timing research. Mol Psychiatry. 2010 Jul;15(7):676-80. PMID: 20571529; X-1, X-4
256. Volker MA, Lopata C, Smerbeck AM, et al. BASC-2 PRS profiles for students with high-functioning autism spectrum disorders. J Autism Dev Disord. 2010 Feb;40(2):188-99. PMID: 19705267; X-1, X-3, X-4
257. Ward SL, Gilmore L. The autistic behavioural indicators instrument (ABII): Development and instrument utility in discriminating autistic disorder from speech and language impairment and typical development. Research in Autism Spectrum Disorders. 2010;4(1):28-42. X-4
258. Absoud M, Parr JR, Salt A, et al. Developing a schedule to identify social communication difficulties and autism spectrum disorder in young children with visual impairment. Dev Med Child Neurol. 2011 Mar;53(3):285-8. PMID: 21166670; X-3
259. Allely CS, Wilson P. Diagnosing autism spectrum disorders in primary care. Practitioner. 2011 Nov;255(1745):27-30, 3. PMID: 22235552; X-1, X-2, X-3, X-4
260. Allen-Brady K, Cai G, Cannon D, et al. No evidence for IL1RAPL1 involvement in selected high-risk autism pedigrees from the AGRE data set. Autism Res. 2011 Aug;4(4):293-6. PMID: 21491612; X-1, X-3 ,X-4
261. Al-Qabandi M, Gorter JW, Rosenbaum P. Early autism detection: are we ready for routine screening? Pediatrics. 2011 Jul;128(1):e211-7. PMID: 21669896; X-2, X-3, X-4, X-5
262. Bethell CD, Kogan MD, Strickland BB, et al. A national and state profile of leading health problems and health care quality for US children: key insurance disparities and across-state variations. Acad Pediatr. 2011 May-Jun;11(3 Suppl):S22-33. PMID: 21570014; X-1, X-4
263. Bolduc ME, Du Plessis AJ, Sullivan N, et al. Spectrum of neurodevelopmental disabilities in children with cerebellar malformations. Dev Med Child Neurol. 2011 May;53(5):409-16. PMID: 21418200; X-4
264. Bolte S, Westerwald E, Holtmann M, et al. Autistic traits and autism spectrum disorders: the clinical validity of two measures presuming a continuum of social communication skills. J Autism Dev Disord. 2011 Jan;41(1):66-72. PMID: 20422277; X-1, X-3 ,X-4
265. Duquette C, Orders S, Fullarton S, et al. Fighting for Their Rights: Advocacy Experiences of Parents of Children Identified with Intellectual Giftedness. Journal for the Education of the Gifted. 2011;34(3):488-512. X-1, X-3,X-4
266. Esposito G, Venuti P, Bornstein MH. Assessment of distress in young children: A comparison of autistic disorder, developmental delay, and typical development. Research in Autism Spectrum Disorders. 2011;5(4):1510-6. X-1, X-3,X-4
267. Gau SS-F, Lee C-M, Lai M-C, et al. Psychometric properties of the Chinese version of the Social Communication Questionnaire. Research in Autism Spectrum Disorders. 2011;5(2):809-18. X-4, X-5
268. Ghuman JK, Leone SL, Lecavalier L, et al. The screen for social interaction (SSI): a screening measure for autism spectrum disorders in preschoolers. Res Dev Disabil. 2011 Nov-Dec;32(6):2519-29. PMID: 21821391; X-3, X-4
269. Gillis JM, Callahan EH, Romanczyk RG. Assessment of Social Behavior in Children with Autism: The Development of the Behavioral Assessment of Social Interactions in Young Children. Research in Autism Spectrum Disorders. 2011;5(1):351-60. X-3, X-4

270. Guo YQ, Tang Y, Rice C, et al. Validation of the Autism Spectrum Screening Questionnaire, Mandarin Chinese Version (CH-ASSQ) in Beijing, China. Autism. 2011 Nov;15(6):713-27. PMID: 21690211; X-3, X-4, X-5
271. Gura GF, Champagne MT, Blood-Siegfried JE. Autism spectrum disorder screening in primary care. Journal of Developmental and Behavioral Pediatrics. 2011;32(1):48-51. X-4
272. Horovitz M, Matson JL, Rieske RD, et al. The relationship between race and challenging behaviours in infants and toddlers with autistic disorder and pervasive developmental disorder-not otherwise specified. Dev Neurorehabil. 2011;14(4):208-14. PMID: 21732805; X-1, X-4
273. Karabekiroğlu K, Akbaş S. Identifying and differentiating PDD-NOS: A comparison with autism and ADHD. Yeni Symposium: psikiyatri, nöroloji ve davraniş bilimleri dergisi. 2011;49(3):141-52. X-1, X-4, X-5
274. Kopp S, Gillberg C. The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): an instrument for better capturing the autism phenotype in girls? A preliminary study involving 191 clinical cases and community controls. Res Dev Disabil. 2011 Nov-Dec;32(6):2875-88. PMID: 21664105; X-3, X-4
275. Lung F-W, Chiang T-L, Lin S-J, et al. Autism-risk screening in the first 3 years of life in Taiwan birth cohort pilot study. Research in Autism Spectrum Disorders. 2011;5(4):1385-9. X-4, X-5
276. Luyster RJ, Kuban KC, O'Shea TM, et al. The Modified Checklist for Autism in Toddlers in extremely low gestational age newborns: individual items associated with motor, cognitive, vision and hearing limitations. Paediatr Perinat Epidemiol. 2011 Jul;25(4):366-76. PMID: 21649679; X-4
277. Manning SE, Davin CA, Barfield WD, et al. Early diagnoses of autism spectrum disorders in Massachusetts birth cohorts, 2001-2005. Pediatrics. 2011 Jun;127(6):1043-51. PMID: 21576313; X-1, X-4
278. Matson JL, Boisjoli JA, Hess JA, et al. Comorbid psychopathology factor structure on the Baby and Infant Screen for Children with aUtIsm Traits-Part 2 (BISCUIT-Part 2). Research in Autism Spectrum Disorders. 2011;5(1):426-32. X-1, X-4
279. Matson JL, Kozlowski AM, Neal D, et al. Cutoffs for the Matson Evaluation of Social Skills with Youngsters-II (MESSY-II) for typically developing children and for children diagnosed with autism spectrum disorders. Research in Autism Spectrum Disorders. 2011;5(2):798-802. X-1, X-4
280. Matson JL, Mahan S, Fodstad JC, et al. Effects of symptoms of co-morbid psychopathology on challenging behaviours among infants and toddlers with Autistic Disorder and PDD-NOS as assessed with the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT). Dev Neurorehabil. 2011;14(3):129-39. PMID: 21548853; X-4
281. Matson JL, Wilkins J, Fodstad JC. The validity of the Baby and Infant Screen for Children with aUtIsm Traits: Part 1 (BISCUIT: Part 1). J Autism Dev Disord. 2011 Sep;41(9):1139-46. PMID: 20177764; X-4
282. Mazefsky CA, Anderson R, Conner CM, et al. Child Behavior Checklist scores for school-aged children with autism: Preliminary evidence of patterns suggesting the need for referral. Journal of Psychopathology and Behavioral Assessment. 2011;33(1):31-7. X-3,X-4
283. McLaughlin MR. Speech and language delay in children. Am Fam Physician. 2011 May 15;83(10):1183-8. PMID: 21568252; X-1, X-2, X-3, X-4
284. Mouridsen SE, Hauschild KM. Autism spectrum disorders in siblings of children with a developmental language disorder. Logoped Phoniatr Vocol. 2011 Dec;36(4):145-9. PMID: 21446849; X-1, X-3, X-4
285. Muratori F, Narzisi A, Tancredi R, et al. The CBCL 1.5-5 and the identification of preschoolers with autism in Italy. Epidemiol Psychiatr Sci. 2011 Dec;20(4):329-38. PMID: 22201210; X-4
286. Murray MJ, Mayes SD, Smith LA. Brief Report: Excellent Agreement between Two Brief Autism Scales (Checklist for Autism Spectrum Disorder and Social Responsiveness Scale) Completed Independently by Parents and the Autism Diagnostic Interview-Revised. Journal of Autism and Developmental Disorders. 2011;41(11):1586-90. X-2, X-3, X-4
287. Ooi YP, Rescorla L, Ang RP, et al. Identification of autism spectrum disorders using the Child Behavior Checklist in Singapore. J Autism Dev Disord. 2011 Sep;41(9):1147-56. PMID: 20405192; X-3, X-4
288. Palmer E, Ketteridge C, Parr JR, et al. Autism spectrum disorder diagnostic assessments: improvements since publication of the National Autism Plan for Children. Arch Dis Child. 2011 May;96(5):473-5. PMID: 20522453; X-3, X-4
289. Paula CS, Ribeiro SH, Fombonne E, et al. Brief report: prevalence of pervasive developmental disorder in Brazil: a pilot study. J Autism Dev Disord. 2011 Dec;41(12):1738-42. PMID: 21337063; X-3, X-4 ,X-5
290. Pierce K, Conant D, Hazin R, et al. Preference for geometric patterns early in life as a risk factor for autism. Arch Gen Psychiatry. 2011 Jan;68(1):101-9. PMID: 20819977; X-1
291. Pinto-Martin JA, Levy SE, Feldman JF, et al. Prevalence of autism spectrum disorder in adolescents born weighing <2000 grams. Pediatrics. 2011 Nov;128(5):883-91. PMID: 22007018; X-3, X-4
292. Pool JL, Hourcade JJ. Developmental Screening: A Review of Contemporary Practice. Education and Training in Autism and Developmental Disabilities. 2011;46(2):267-75. X-2, X-3, X-4
293. Rowland C. Using the Communication Matrix to Assess Expressive Skills in Early Communicators. Communication Disorders Quarterly. 2011;32(3):190-201. X-1, X-3, X-4
294. Sipes M, Matson JL, Horovitz M. Autism spectrum disorders and motor skills: the effect on socialization as measured by the Baby And Infant Screen For Children with aUtIsm Traits (BISCUIT). Dev Neurorehabil. 2011;14(5):290-6. PMID: 21870953; X-1, X-4

Appendix E. Excluded Studies

295. Sipes M, Matson JL, Turygin N. The use of the Battelle Developmental Inventory-Second Edition (BDI-2) as an early screener for autism spectrum disorders. Dev Neurorehabil. 2011;14(5):310-4. PMID: 21870955; X-3
296. Suzuki M, Tachimori H, Saito M, et al. Development of a screening scale for high-functioning pervasive developmental disorders using the Tokyo Child Development Schedule and Tokyo Autistic Behavior Scale. Research in Autism Spectrum Disorders. 2011;5(2):843-54. X-3, X-4
297. Tonge B, Brereton A. Autism spectrum disorders. Aust Fam Physician. 2011 Sep;40(9):672-7. PMID: 21894273; X-2, X-3 ,X-4
298. van Daalen E, Kemner C, Verbeek NE, et al. Social Responsiveness Scale-aided analysis of the clinical impact of copy number variations in autism. Neurogenetics. 2011 Nov;12(4):315-23. PMID: 21837366; X-3, X-4
299. Vanvuchelen M, Roeyers H, De Weerdt W. Imitation Assessment and Its Utility to the Diagnosis of Autism: Evidence from Consecutive Clinical Preschool Referrals for Suspected Autism. Journal of Autism and Developmental Disorders. 2011;41(4):484-96. X-1, X-4
300. Webb PL. Screening for Autism Spectrum Disorders During Well-Child Visits in a Primary Care Setting. Journal for Nurse Practitioners. 2011;7(3):229-35. X-4
301. Worley JA, Matson JL, Kozlowski AM. The effects of hearing impairment on symptoms of autism in toddlers. Dev Neurorehabil. 2011;14(3):171-6. PMID: 21548858; X-4
302. Worley JA, Matson JL, Mahan S, et al. Stability of symptoms of autism spectrum disorders in toddlers: an examination using the Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT). Dev Neurorehabil. 2011;14(1):36-40. PMID: 21241177; X-4
303. Worley JA, Matson JL, Sipes M, et al. Prevalence of autism spectrum disorders in toddlers receiving early intervention services. Research in Autism Spectrum Disorders. 2011;5(2):920-5. X-4
304. Yamada Y, Yoshida F, Hemmi H, et al. Atypical social development in neonatal intensive care unit survivors at 12 months. Pediatr Int. 2011 Dec;53(6):858-66. PMID: 21429057; X-4
305. Young GS, Rogers SJ, Hutman T, et al. Imitation from 12 to 24 months in autism and typical development: A longitudinal Rasch analysis. Developmental Psychology. 2011;47(6):1565-78. X-1, X-4
306. Allison C, Auyeung B, Baron-Cohen S. Toward brief "Red Flags" for autism screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist for Autism in toddlers in 1,000 cases and 3,000 controls [corrected]. J Am Acad Child Adolesc Psychiatry. 2012 Feb;51(2):202-12 e7. PMID: 22265366; X-4
307. Arunyanart W, Fenick A, Ukritchon S, et al. Developmental and Autism Screening: A Survey Across Six States. Infants & Young Children: An Interdisciplinary Journal of Special Care Practices. 2012;25(3):175-87. X-3, X-4
308. Aydin A. Turkish Adaptation of Test of Pretended Play. Educational Sciences: Theory and Practice. 2012;12(2):916-25. X-1, X-3, X-4, X-5
309. Baghdadli A, Assouline B, Sonie S, et al. Developmental trajectories of adaptive behaviors from early childhood to adolescence in a cohort of 152 children with autism spectrum disorders. J Autism Dev Disord. 2012 Jul;42(7):1314-25. PMID: 21928042; X-1, X-3 ,X-4
310. Barton ML, Dumont-Mathieu T, Fein D. Screening Young Children for Autism Spectrum Disorders in Primary Practice. Journal of Autism and Developmental Disorders. 2012;42(6):1165-74. X-2, X-3, X-4
311. Ben-Sasson A, Carter AS. The application of the First Year Inventory for ASD screening in Israel. Journal of Autism and Developmental Disorders. 2012;42(9):1906-16. X-4
312. Bolton PF, Golding J, Emond A, et al. Autism spectrum disorder and autistic traits in the Avon Longitudinal Study of Parents and Children: precursors and early signs. J Am Acad Child Adolesc Psychiatry. 2012 Mar;51(3):249-60 e25. PMID: 22365461; X-1, X-4
313. Coo H, Ouellette-Kuntz H, Lam M, et al. Correlates of age at diagnosis of autism spectrum disorders in six Canadian regions. Chronic Dis Inj Can. 2012 Mar;32(2):90-100. PMID: 22414306; X-1, X-3, X-4
314. Dereu M, Raymaekers R, Warreyn P, et al. Can child care workers contribute to the early detection of autism spectrum disorders? A comparison between screening instruments with child care workers versus parents as informants. J Autism Dev Disord. 2012 May;42(5):781-96. PMID: 21691866; X-1
315. Dereu M, Roeyers H, Raymaekers R, et al. How useful are screening instruments for toddlers to predict outcome at age 4? General development, language skills, and symptom severity in children with a false positive screen for autism spectrum disorder. Eur Child Adolesc Psychiatry. 2012 Oct;21(10):541-51. PMID: 22580987; X-1
316. DeWeerdt S. Culture: Diverse diagnostics. Nature. 2012 Nov 1;491(7422):S18-9. PMID: 23136656; X-1, X-2, X-3, X-4, X-5
317. Feldman MA, Ward RA, Savona D, et al. Development and initial validation of a parent report measure of the behavioral development of infants at risk for autism spectrum disorders. J Autism Dev Disord. 2012 Jan;42(1):13-22. PMID: 21387111;
318. Fisher B, Dezort C, Nordli DR, et al. Routine developmental and autism screening in an epilepsy care setting. Epilepsy Behav. 2012 Aug;24(4):488-92. PMID: 22789633; X-4
319. Fombonne E, Marcin C, Bruno R, et al. Screening for autism in Mexico. Autism Res. 2012 Jun;5(3):180-9. PMID: 22581514; X-3, X-4 ,X-5
320. Frazier TW, Youngstrom EA, Speer L, et al. Validation of proposed DSM-5 criteria for autism spectrum disorder. J Am Acad Child Adolesc Psychiatry. 2012 Jan;51(1):28-40 e3. PMID: 22176937; X-3, X-4
321. Grinker RR, Chambers N, Njongwe N, et al. "Communities" in community engagement: lessons learned from autism research in South Korea and South Africa. Autism Res. 2012 Jun;5(3):201-10. PMID: 22566396; X-4, X-5

322. Hattier MA, Matson JL. An examination of the relationship between communication and socialization deficits in children with autism and PDD-NOS. Research in Autism Spectrum Disorders. 2012;6(2):871-80. X-1, X-4
323. Hattier MA, Matson JL, Kozlowski AM. Communication skills in children with cerebral palsy and autism spectrum disorder. Journal of Developmental and Physical Disabilities. 2012;24(1):85-93. X-1, X-3, X-4
324. Hattier MA, Matson JL, May AC, et al. Repetitive/restricted behaviours and interests in children with cerebral palsy and autism spectrum disorder. Dev Neurorehabil. 2012;15(3):178-84. PMID: 22582848; X-1, X-3, X-4
325. Honigfeld L, Chandhok L, Spiegelman K. Engaging pediatricians in developmental screening: the effectiveness of academic detailing. J Autism Dev Disord. 2012 Jun;42(6):1175-82. PMID: 21826585; X-4
326. Hutman T, Chela MK, Gillespie-Lynch K, et al. Selective visual attention at twelve months: signs of autism in early social interactions. J Autism Dev Disord. 2012 Apr;42(4):487-98. PMID: 21519953; X-4
327. Hyman SL, Johnson JK. Autism and Pediatric Practice: Toward a Medical Home. Journal of Autism and Developmental Disorders. 2012;42(6):1156-64. X-2, X-3 ,X-4
328. Ito H, Tani I, Yukihiro R, et al. Validation of an interview-based rating scale developed in Japan for pervasive developmental disorders. Research in Autism Spectrum Disorders. 2012;6(4):1265-72. X-1, X-3 ,X-4
329. Kim SH, Lord C. New Autism Diagnostic Interview-Revised Algorithms for Toddlers and Young Preschoolers from 12 to 47 Months of Age. Journal of Autism and Developmental Disorders. 2012;42(1):82-93. X-1, X-4
330. Kozlowski AM, Matson JL, Worley JA, et al. Defining characteristics for young children meeting cutoff on the Modified Checklist for Autism in Toddlers. Research in Autism Spectrum Disorders. 2012;6(1):472-9.
331. Lakic A. Screening, identification and evaluation of autism spectrum disorders in primary health care. Vojnosanit Pregl. 2012 May;69(5):437-43. PMID: 22764548; X-2, X-3, X-4, X-5
332. Lialiou P, Zikos D, Mantas J. Development and evaluation of an expert system for the diagnosis of child autism. Stud Health Technol Inform. 2012;180:1185-7. PMID: 22874395; X-3, X-4
333. Lord C, Luyster R, Guthrie W, et al. Patterns of developmental trajectories in toddlers with autism spectrum disorder. J Consult Clin Psychol. 2012 Jun;80(3):477-89. PMID: 22506796; X-1, X-3, X-4
334. Losh M, Esserman D, Anckarsater H, et al. Lower birth weight indicates higher risk of autistic traits in discordant twin pairs. Psychol Med. 2012 May;42(5):1091-102. PMID: 22132806; X-1, X-3, X-4
335. LoVullo SV, Matson JL. Development of a critical item algorithm for the Baby and Infant Screen for Children with aUtIsm Traits. Research in Autism Spectrum Disorders. 2012;6(1):378-84.
336. Magyar CI, Pandolfi V, Dill CA. An initial evaluation of the Social Communication Questionnaire for the assessment of autism spectrum disorders in children with Down syndrome. J Dev Behav Pediatr. 2012 Feb;33(2):134-45. PMID: 22267105; X-3, X-4
337. Mandell DS, Lawer LJ, Branch K, et al. Prevalence and Correlates of Autism in a State Psychiatric Hospital. Autism: The International Journal of Research and Practice. 2012;16(6):557-67. X-3,X-4, X-5
338. Matson JL, Kozlowski AM, Hattier MA, et al. DSM-IV vs DSM-5 diagnostic criteria for toddlers with autism. Dev Neurorehabil. 2012;15(3):185-90. PMID: 22582849; X-4
339. Matson JL, Tureck K. Early Diagnosis of Autism: Current Status of the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT-Parts 1, 2, and 3). Research in Autism Spectrum Disorders. 2012;6(3):1135-41. X-1, X-2, X-3, X-4, X-5
340. Medeiros K, Kozlowski AM, Beighley JS, et al. The effects of developmental quotient and diagnostic criteria on challenging behaviors in toddlers with developmental disabilities. Res Dev Disabil. 2012 Jul-Aug;33(4):1110-6. PMID: 22502836; X-1, X-3 ,X-4
341. Moore T, Johnson S, Hennessy E, et al. Screening for autism in extremely preterm infants: problems in interpretation. Dev Med Child Neurol. 2012 Jun;54(6):514-20. PMID: 22458327; X-4
342. Nakanishi M, Deardorff MA, Clark D, et al. Investigation of autistic features among individuals with mild to moderate Cornelia de Lange syndrome. Am J Med Genet A. 2012 Aug;158A(8):1841-7. PMID: 22740374; X-3, X-4
343. Nath S, Roy R, Mukherjee S. Perinatal complications associated with autism--a case control study in a neurodevelopment and early intervention clinic. J Indian Med Assoc. 2012 Aug;110(8):526-9. PMID: 23741815; X-4, X-5
344. Neal D, Matson JL, Belva BC. Discriminant analysis of the autism spectrum disorder observation for children. Developmental Neurorehabilitation. 2012;15(4):267-73. X-3, X-4
345. Nygren G, Cederlund M, Sandberg E, et al. The prevalence of autism spectrum disorders in toddlers: a population study of 2-year-old Swedish children. J Autism Dev Disord. 2012 Jul;42(7):1491-7. PMID: 22048962; X-4
346. Peacock G, Lin SC. Enhancing early identification and coordination of intervention services for young children with autism spectrum disorders: report from the Act Early Regional Summit Project. Disabil Health J. 2012 Jan;5(1):55-9. PMID: 22226299; X-1, X-2, X-3, X-4
347. Quek LH, Sofronoff K, Sheffield J, et al. Co-occurring anger in young people with Asperger's syndrome. Journal of Clinical Psychology. 2012;68(10):1142-8. X-1, X-3, X-4
348. Reel KH, Lecavalier L, Butter E, et al. Diagnostic utility of the Pervasive Developmental Disorder Behavior Inventory. Research in Autism Spectrum Disorders. 2012;6(1):458-65. X-3, X-4

Appendix E. Excluded Studies

349. Ross JL, Roeltgen DP, Kushner H, et al. Behavioral and social phenotypes in boys with 47,XYY syndrome or 47,XXY Klinefelter syndrome. Pediatrics. 2012 Apr;129(4):769-78. PMID: 22412026; X-3, X-4
350. Roux AM, Herrera P, Wold CM, et al. Developmental and autism screening through 2-1-1: reaching underserved families. Am J Prev Med. 2012 Dec;43(6 Suppl 5):S457-63. PMID: 23157765; X-1, X-3, X-4
351. Sasson NJ, Elison JT. Eye tracking young children with autism. J Vis Exp. 2012(61)PMID: 22491039; X-1, X-2 , X-3, X-4
352. Schieve LA, Boulet SL, Blumberg SJ, et al. Association between parental nativity and autism spectrum disorder among US-born non-Hispanic white and Hispanic children, 2007 National Survey of Children's Health. Disabil Health J. 2012 Jan;5(1):18-25. PMID: 22226294; X-1, X-3, X-4
353. Soares NS, Patel DR. Office screening and early identification of children with autism. Pediatr Clin North Am. 2012 Feb;59(1):89-102, x-xi. PMID: 22284795; X-2, X-3, X-4, X-5
354. Stephens BE, Bann CM, Watson VE, et al. Screening for autism spectrum disorders in extremely preterm infants. J Dev Behav Pediatr. 2012 Sep;33(7):535-41. PMID: 22926660; X-4
355. Tsai WC, Soong WT, Shyu YI. Toddler autism screening questionnaire: development and potential clinical validity. Autism. 2012 Jul;16(4):340-9. PMID: 22399452; X-5
356. Valicenti-McDermott M, Hottinger K, Seijo R, et al. Age at diagnosis of autism spectrum disorders. J Pediatr. 2012 Sep;161(3):554-6. PMID: 22683037; X-1, X-4
357. Veness C, Prior M, Bavin E, et al. Early indicators of autism spectrum disorders at 12 and 24 months of age: a prospective, longitudinal comparative study. Autism. 2012 Mar;16(2):163-77. PMID: 21733958; X-4
358. Wall DP, Dally R, Luyster R, et al. Use of artificial intelligence to shorten the behavioral diagnosis of autism. PLoS ONE. 2012;7(8) X-1, X-3, X-4
359. Wang J, Lee L-C, Chen Y-S, et al. Assessing autistic traits in a Taiwan preschool population: Cross-cultural validation of the Social Responsiveness Scale (SRS). Journal of Autism and Developmental Disorders. 2012;42(11):2450-9. X-3, X-4, X-5
360. Warren Z, Vehorn A, Dohrmann E, et al. Accuracy of phenotyping children with autism based on parent report: what specifically do we gain phenotyping "rapidly"? Autism Res. 2012 Feb;5(1):31-8. PMID: 21972233; X-3, X-4
361. Wegiel J, Schanen NC, Cook EH, et al. Differences between the pattern of developmental abnormalities in autism associated with duplications 15q11.2-q13 and idiopathic autism. J Neuropathol Exp Neurol. 2012 May;71(5):382-97. PMID: 22487857; X-1 ,X-3 ,X-4
362. Wiggins LD, Robins DL, Adamson LB, et al. Support for a dimensional view of autism spectrum disorders in toddlers. J Autism Dev Disord. 2012 Feb;42(2):191-200. PMID: 21448751; X-1, X-4
363. Wright K, Poulin-Dubois D. Modified Checklist for Autism in Toddlers (M-CHAT) Screening at 18 Months of Age Predicts Concurrent Understanding of Desires, Word Learning and Expressive Vocabulary. Research in Autism Spectrum Disorders. 2012;6(1):184-92. X-4
364. Yama B, Freeman T, Graves E, et al. Examination of the properties of the Modified Checklist for Autism in Toddlers (M-CHAT) in a population sample. J Autism Dev Disord. 2012 Jan;42(1):23-34. PMID: 21373956; X-4
365. Zwaigenbaum L, Bryson SE, Szatmari P, et al. Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. J Autism Dev Disord. 2012 Dec;42(12):2585-96. PMID: 22453928; X-1, X-4
366. Andersson GW, Gillberg C, Miniscalco C. Pre-school children with suspected autism spectrum disorders: Do girls and boys have the same profiles? Research in Developmental Disabilities. 2013;34(1):413-22. X-1, X-4
367. Barbaro J, Dissanayake C. Early markers of autism spectrum disorders in infants and toddlers prospectively identified in the Social Attention and Communication Study. Autism. 2013 Jan;17(1):64-86. PMID: 22735682; X-4
368. Barton ML, Robins DL, Jashar D, et al. Sensitivity and Specificity of Proposed DSM-5 Criteria for Autism Spectrum Disorder in Toddlers. Journal of Autism and Developmental Disorders. 2013 May 2013;43(5):1184-95. X-1, X-3, X-4
369. Begeer S, Mandell D, Wijnker-Holmes B, et al. Sex Differences in the Timing of Identification Among Children and Adults with Autism Spectrum Disorders. Journal of Autism and Developmental Disorders. 2013;43(5):1151-6. X-1, X-3, X-4
370. Ben-Sasson A, Amit-Ben-Simhon H, Meyer S. Cross-parent reliability in rating ASD markers in infants. Dev Neurorehabil. 2013 Aug 7PMID: 23924084; X-4
371. Ben-Sasson A, Carter AS. The contribution of sensory–regulatory markers to the accuracy of ASD screening at 12 months. Research in Autism Spectrum Disorders. 2013;7(7):879-88. X-4
372. Beuker KT, Schjølberg S, Lie KK, et al. The structure of autism spectrum disorder symptoms in the general population at 18 months. Journal of Autism and Developmental Disorders. 2013;43(1):45-56. X-1, X3, X-4
373. Bond S. EPIDEMIOLOGIC STUDY LINKS INDUCTION AND AUGMENTATION OF LABOR TO DIAGNOSIS OF AUTISM. J Midwifery Womens Health. 2013 Dec 10PMID: 24325141; X-1, X-2, X-3, X-4
374. Charman T, Gotham K. Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practice. Child Adolesc Ment Health. 2013 Feb 1;18(1):52-63. PMID: 23539140; X-2, X-3, X-4
375. Chiang C-H, Wu C-C, Hou Y-M, et al. Development of T-STAT for Early Autism Screening. Journal of Autism and Developmental Disorders. 2013;43(5):1028-37. X-5

376. Damiano CR, Nahmias A, Hogan-Brown AL, et al. What do repetitive and stereotyped movements mean for infant siblings of children with autism spectrum disorders? Journal of Autism and Developmental Disorders. 2013;43(6):1326-35. X-1, X-4
377. Daniels AM, Mandell DS. Children's Compliance with American Academy of Pediatrics' Well-Child Care Visit Guidelines and the Early Detection of Autism. Journal of Autism & Developmental Disorders. 2013;43(12):2844-54. X-1, X-3, X-4
378. Davidson C, Greenwood N, Stansfield A, et al. Prevalence of Asperger syndrome among patients of an Early Intervention in Psychosis team. Early Interv Psychiatry. 2013 Mar 8PMID: 23472601; X-3, X-4
379. Elmose M, Trillingsgaard A, Jorgensen M, et al. Follow-up at mid-school age (9-13 years) of children assessed for autism spectrum disorder before the age of four. Nord J Psychiatry. 2013 Nov 7PMID: 24199947; X-1, X-3, X-4
380. Esposito G, Del Carmen Rostagno M, Venuti P, et al. Brief Report: Atypical Expression of Distress During the Separation Phase of the Strange Situation Procedure in Infant Siblings at High Risk for ASD. J Autism Dev Disord. 2013 Sep 12PMID: 24026913; X-1, X-3, X-4
381. Frazier TW, Youngstrom EA, Embacher R, et al. Demographic and clinical correlates of autism symptom domains and autism spectrum diagnosis. Autism. 2013 Oct 8PMID: 24104512; X-1, X-3, X-4
382. Freuler AC, Baranek GT, Tashjian C, et al. Parent reflections of experiences of participating in a randomized controlled trial of a behavioral intervention for infants at risk of autism spectrum disorders. Autism. 2013 Oct 8PMID: 24104508; X-1, X-4
383. Gan SM, Tung LC, Yeh CH, et al. The ICF-CY-based structural equation model of factors associated with participation in children with autism. Dev Neurorehabil. 2013 Oct 2PMID: 24087923; X-1, X-3, X-4, X-5
384. Gardner LM, Murphy L, Campbell JM, et al. Screening accuracy for risk of autism spectrum disorder using the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). Research in Autism Spectrum Disorders. 2013;7(5):591-600. X-4
385. Garg S, Green J, Leadbitter K, et al. Neurofibromatosis type 1 and autism spectrum disorder. Pediatrics. 2013 Dec;132(6):e1642-8. PMID: 24190681; X-3, X-4
386. Garg S, Lehtonen A, Huson SM, et al. Autism and other psychiatric comorbidity in neurofibromatosis type 1: evidence from a population-based study. Dev Med Child Neurol. 2013 Feb;55(2):139-45. PMID: 23163236; X-3, X-4
387. Gau SS-F, Liu L-T, Wu Y-Y, et al. Psychometric properties of the Chinese version of the Social Responsiveness Scale. Research in Autism Spectrum Disorders. 2013;7(2):349-60. X-3, X-4, X-5
388. Guevara JP, Gerdes M, Localio R, et al. Effectiveness of developmental screening in an urban setting. Pediatrics. 2013 Jan;131(1):30-7. PMID: 23248223; X-1, X-3, X-4
389. Harrington JW, Bai R, Perkins AM. Screening children for autism in an urban clinic using an electronic M-CHAT. Clin Pediatr (Phila). 2013 Jan;52(1):35-41. PMID: 23117237; X-4
390. Harris B, Barton EE, Albert C. Evaluating Autism Diagnostic and Screening Tools for Cultural and Linguistic Responsiveness. J Autism Dev Disord. 2013 Nov 2PMID: 24186120; X-1, X-3, X-4
391. Hedvall A, Westerlund J, Fernell E, et al. Autism and developmental profiles in preschoolers: stability and change over time. Acta Paediatr. 2013 Oct 8PMID: 24237479; X-1, X-3, X-4
392. Horovitz M, Matson JL. The baby and infant screen for children with autism traits-part 2: The development of age-based cutoffs. Dev Neurorehabil. 2013 Jul 19PMID: 23869771; X-1, X-4
393. Hu VW, Lai Y. Developing a Predictive Gene Classifier for Autism Spectrum Disorders Based upon Differential Gene Expression Profiles of Phenotypic Subgroups. N Am J Med Sci (Boston). 2013;6(3)PMID: 24363828; X-1, X-3, X-4
394. Hus V, Bishop S, Gotham K, et al. Factors influencing scores on the Social Responsiveness Scale. Journal of Child Psychology and Psychiatry. 2013;54(2):216-24. X-3, X-4
395. Jain R, Juneja M, Sairam S. Children with developmental disabilities in India: Age of initial concern and referral for rehabilitation services, and reasons for delay in referral. Journal of Child Neurology. 2013;28(4):455-60. X-1, X-3, X-4, X-5
396. Jaspers M, de Winter AF, Buitelaar JK, et al. Early childhood assessments of community pediatric professionals predict autism spectrum and attention deficit hyperactivity problems. J Abnorm Child Psychol. 2013 Jan;41(1):71-80. PMID: 22688680; X-1, X-3, X-4
397. Jutley-Neilson J, Harris G, Kirk J. The identification and measurement of autistic features in children with septo-optic dysplasia, optic nerve hypoplasia and isolated hypopituitarism. Res Dev Disabil. 2013 Dec;34(12):4310-8. PMID: 24210356; X-3, X-4
398. Kakooza-Mwesige A, Ssebyala K, Karamagi C, et al. Adaptation of the 'ten questions' to screen for autism and other neuro-developmental disorders in Uganda. Autism. 2013 Nov 8PMID: 23536263; X-4, X-5
399. Kantzer AK, Fernell E, Gillberg C, et al. Autism in community pre-schoolers: developmental profiles. Res Dev Disabil. 2013 Sep;34(9):2900-8. PMID: 23816626; X-1, X-4
400. Kirby A, Sugden D, Purcell C. Diagnosing developmental coordination disorders. Arch Dis Child. 2013 Nov 19PMID: 24255567; X-1, X-2, X-3, X-4
401. Koh HC, Lim SH, Chan GJ, et al. The Clinical Utility of the Modified Checklist for Autism in Toddlers with High Risk 18-48 Month Old Children in Singapore. J Autism Dev Disord. 2013 Jun 28PMID: 23812662;
402. Kumar VS, Webster M. A test in development. Clin Chem. 2013 Nov;59(11):1680-2. PMID: 24167190; X-1, X-2, X-3, X-4
403. Liu X, Zhou X, Lackaff J. Incremental Validity in the Clinical Assessment of Early Childhood Development. Journal of Psychoeducational Assessment. 2013 Oct 2013;31(5):506-15. X-1, X-3, X-4
404. Lukose R, Brown K, Barber CM, et al. Quantification of the stapedial reflex reveals delayed responses in autism. Autism Res. 2013 Oct;6(5):344-53. PMID: 23825093; X-1, X-3, X-4

405. Matson JL, Kozlowski AM, Fitzgerald ME, et al. True versus false positives and negatives on the Modified Checklist For Autism in Toddlers. Research in Autism Spectrum Disorders. 2013;7(1):17-22. X-4
406. Mayes SD, Black A, Tierney CD. DSM-5 under-identifies PDDNOS: Diagnostic agreement between the DSM-5, DSM-IV, and Checklist for Autism Spectrum Disorder. Research in Autism Spectrum Disorders. 2013;7(2):298-306. X-1, X-3, X-4
407. Myers CL, Gross AD, McReynolds BM. Broadband Behavior Rating Scales as Screeners for Autism? J Autism Dev Disord. 2013 Nov 28PMID: 24287879;
408. Nakai Y, Takashima R, Takiguchi T, et al. Speech intonation in children with autism spectrum disorder. Brain Dev. 2013 Aug 20PMID: 23973369; X-1, X-3, X-4
409. Narzisi A, Calderoni S, Maestro S, et al. Child Behavior Check List 1(1/2)-5 as a tool to identify toddlers with autism spectrum disorders: a case-control study. Res Dev Disabil. 2013 Apr;34(4):1179-89. PMID: 23376628; X-4
410. Neal D, Matson JL, Belva BC. An examination of the reliability of a new observation measure for Autism spectrum disorders: The autism spectrum disorder observation for children. Research in Autism Spectrum Disorders. 2013;7(1):29-34. X-1, X-3, X-4
411. Nickel LR, Thatcher AR, Keller F, et al. Posture Development in Infants at Heightened vs. Low Risk for Autism Spectrum Disorders. Infancy. 2013 Sep;18(5):639-61. PMID: 24027437; X-1, X-3, X-4
412. Nordenbaek C, Jorgensen M, Kyvik KO, et al. A Danish population-based twin study on autism spectrum disorders. Eur Child Adolesc Psychiatry. 2013 May 10PMID: 23661220; X-3, X-4
413. Oner P, Oner O, Munir K. Three-item Direct Observation Screen (TIDOS) for autism spectrum disorder. Autism. 2013 Oct 14PMID: 24126869; X-4, X-5
414. Pringsheim T, Hammer T. Social behavior and comorbidity in children with tics. Pediatr Neurol. 2013 Dec;49(6):406-10. PMID: 24095577; X-1, X-3, X-4
415. Robinson EB, Howrigan D, Yang J, et al. Response to 'Predicting the diagnosis of autism spectrum disorder using gene pathway analysis'. Mol Psychiatry. 2013 Oct 22PMID: 24145379; X-1, X-3, X-4
416. Samadi SA, McConkey R. The utility of the Gilliam autism rating scale for identifying Iranian children with autism. Disabil Rehabil. 2013 Jun 5PMID: 23738615; X-3, X-4, X-5
417. Sasson NJ, Lam KSL, Childress D, et al. The Broad Autism Phenotype Questionnaire: Prevalence and diagnostic classification. Autism Research. 2013;6(2):134-43. X-1, X-3, X-4
418. Scarpa A, Reyes NM, Patriquin MA, et al. The Modified Checklist for Autism in Toddlers: Reliability in a diverse rural American sample. Journal of Autism and Developmental Disorders. 2013;43(10):2269-79. X-4
419. Shibata A, Hitomi Y, Kambayashi Y, et al. Epidemiological study on the involvements of environmental factors and allergy in child mental health using the Autism Screening Questionnaire. Research in Autism Spectrum Disorders. 2013;7(1):132-40. X-3, X-4
420. Shrestha M, Shrestha R. Symptom Recognition to Diagnosis of Autism in Nepal. J Autism Dev Disord. 2013 Nov 29PMID: 24293038; X-1, X-3, X-4, X-5
421. Sipes M, Matson JL. Factor Structure for Autism Spectrum Disorders with Toddlers Using DSM-IV and DSM-5 Criteria. J Autism Dev Disord. 2013 Aug 25PMID: 23979069; X-1, X-4
422. Smith NJ, Sheldrick RC, Perrin EC. An abbreviated screening instrument for autism spectrum disorders. Infant Mental Health Journal. 2013;34(2):149-55. X-4
423. So P, Greaves-Lord K, van der Ende J, et al. Using the Child Behavior Checklist and the Teacher's Report Form for identification of children with autism spectrum disorders. Autism. 2013;17(5):595-607. X-3, X-4
424. Sun X, Allison C, Auyeung B, et al. The Mandarin Chinese version of the Childhood Autism Spectrum Test (CAST): Test–retest reliability. Research in Developmental Disabilities. 2013;34(10):3267-75. X-3, X-4, X-5
425. Sun X, Allison C, Auyeung B, et al. The Mandarin Chinese version of the childhood autism spectrum test (CAST): test-retest reliability. Res Dev Disabil. 2013 Oct;34(10):3267-75. PMID: 23892873; X-3, X-4, X-5
426. Sun X, Allison C, Matthews FE, et al. Prevalence of autism in mainland China, Hong Kong and Taiwan: a systematic review and meta-analysis. Mol Autism. 2013;4(1):7. PMID: 23570419; X-1, X-2, X-3, X-4, X-5
427. Tang LL, Kang L, Liu C, et al. An additive selection of markers to improve diagnostic accuracy based on a discriminatory measure. Acad Radiol. 2013 Jul;20(7):854-62. PMID: 23611438; X-1, X-3, X-4
428. Taylor LJ, Maybery MT, Whitehouse AJ. Moving beyond behaviour-only assessment: Incorporating biomarkers to improve the early detection and diagnosis of autism spectrum disorders. Int J Speech Lang Pathol. 2013 Nov 18PMID: 24236911; X-1, X-2, X-3, X-4
429. Tomlin A, Koch SM, Raches C, et al. Autism Screening Practices Among Early Intervention Providers in Indiana. Infants & Young Children: An Interdisciplinary Journal of Special Care Practices. 2013 2013 Jan-Mar;26(1):74-88. X-1, X-3, X-4
430. Travers JC, Tincani M, Krezmien MP. A Multiyear National Profile of Racial Disparity in Autism Identification. Journal of Special Education. 2013;47(1):41-9. X-1, X-3, X-4
431. Tuominen-Eriksson A-M, Svensson Y, Gunnarsson RK. Children with Disabilities Are Often Misdiagnosed Initially and Children with Neuropsychiatric Disorders Are Referred to Adequate Resources 30 Months Later than Children with Other Disabilities. Journal of Autism and Developmental Disorders. 2013;43(3):579-84. X-1, X-3, X-4
432. Volkmar FR, Reichow B. Infants and toddlers with autism: The promise and the challenges. Int J Speech Lang Pathol. 2013 Dec 18PMID: 24345000; X-1, X-2, X-3, X-4

Appendix E. Excluded Studies

433. Wiggins LD, Robins DL, Yeargin-Allsopp M. Short report: Improving record-review surveillance of young children with an autism spectrum disorder. Autism. 2013 Sep;17(5):623-9. PMID: 22987886; X-1, X-3, X-4
434. Williams K, Perkins D, Wheeler D, et al. Can questions about social interaction correctly identify preschool aged children with autism? J Paediatr Child Health. 2013 Feb;49(2):E167-74. PMID: 23350819;
435. Wong HS, Huertas-Ceballos A, Cowan FM, et al. Evaluation of early childhood social-communication difficulties in children born preterm using the Quantitative Checklist for Autism in Toddlers. J Pediatr. 2014 Jan;164(1):26-33.e1. PMID: 23972644; X-4

Appendix F Table 1. Summary of Adaptive Behavior Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
RCTs					
Dawson et al, 2012[1, 2] US **IG:** ESDM, 24/24 **CG:** Eclectic therapy, 24/21 **Quality:** Good	**IG:** 23.9 ± 4.0 **CG:** 23.1 ± 3.9	VABS Socialization	**IG:** 73.8 ± 7.7 **CG:** 72.4 ± 9.4 **IG/CG:** *P* = 0.594	At 2 years: **IG:** 69.2 ± 11.6 **CG:** 63.1 ± 9.3 **IG/CG:** *P* = 0.263	No baseline group differences in any of the VABS domain scores Significant group x time differences in all reported VABS domain scores except socialization
		Communication	**IG:** 68.4 ± 7.6 **CG:** 69.6 ± 7.3 **IG/CG:** *P* = 0.577	**IG:** 82.1 ± 21.8 **CG:** 69.4 ± 15.8 **IG/CG:** *P* = 0.015	
		Composite score	**IG:** 69.5 ± 5.7 **CG:** 69.9 ± 7.3 **IG/CG:** *P* = 0.844	**IG:** 68.7 ± 15.9 **CG:** 59.1 ± 8.8 **IG/CG:** *P* = 0.011	
		Daily living skills	**IG:** 87.3 ± 11.4 **CG:** 86.8 ± 10.0 **IG/CG:** *P* = 0.381	**IG:** 64.7 ± 12.4 **CG:** 58.0 ± 8.1 **IG/CG:** *P* = 0.013	
		Motor skills	**IG:** 70.9 ± 6.2 **CG:** 72.5 ± 6.5 **IG/CG:** *P* = 0.862	**IG:** 77.4 ± 19.8 **CG:** 64.1 ± 12.3 **IG/CG:** *P* = 0.009	
Sallows et al, 2005[3] **IG:** Clinic directed UCLA/Lovaas-based early intensive intervention, 13/13 **CG:** Parent-directed UCLA/Lovaas-based early intensive intervention, 10/10 **Ga:** Rapid learners by Early Learning Measure **Gb:** Moderate learners by Early Learning Measure **Quality:** Good	**IG:** 33.23 ± 3.89 **CG:** 34.20 ± 5.06	VABS Communication	**IG:** 57.46 ± 4.97 **CG:** 63.20 ± 5.58 **Ga:** 60.82 ± 4.02 **Gb:** 59.17 ± 7.22	**IG:** 73.69 ± 32.32 **CG:** 81.40 ± 24.33 **Ga:** 105.09 ± 12.83 **Gb:** 51.33 ± 10.94 Teacher rated: **Ga:** 94.44 ± 13.97 **Gb:** 58.58 ± 7.90	No significant differences between groups (IG/CG) at pre- or posttest. Combining children in both groups (IG+ CG), pretest to posttest gains were significant for Vineland Communication, F(1, 21)=7.57, p<0.05, Vineland Socialization, F(1, 21)=10.30, p<0.01 Ga showed significant gains in all areas measured: Communication, F(1, 21)= 147.07, p<0.01; Daily Living Skills, F(1, 21)=20.50, p<0.01; Socialization, F(1, 21)=42.89, p<0.01; Applied Behavior Composite, F(1, 21)=54.17, p<0.01 Significant group differences (Ga vs. Gb) in the Teacher ratings of Communication (t=6.84, p<0.01) and Socialization (t=4.60, p<0.01)
		Daily living skills	**IG:** 63.92 ± 5.53 **CG:** 64.20 ± 3.68 **Ga:** 66.45 ± 4.25 **Gb:** 61.83 ± 4.20	**IG:** 66.23 ± 25.95 **CG:** 64.20 ± 12.42 **Ga:** 82.27 ± 16.34 **Gb:** 49.83 ± 10.61	
		Composite	**IG:** 59.54 ± 5.31 **CG:** 60.90 ± 5.94 **Ga:** 61.73 ± 4.59 **Gb:** 58.67 ± 6.09	**IG:** 69.00 ± 28.04 **CG:** 66.70 ± 14.68 **Ga:** 88.64 ± 15.68 **Gb:** 49.08 ± 7.76	
		Socialization	**IG:** 58.38 ± 6.17 **CG:** 60.30 ± 5.76 **Ga:** 61.55 ± 6.58 **Gb:** 57.08 ± 4.63	**IG:** 73.92 ± 23.49 **CG:** 68.90 ± 10.11 **Ga:** 87.73 ± 14.94 **Gb:** 57.08 ± 6.40 Socialization teacher rated: **Ga:** 89.89 ± 18.36 **Gb:** 61.58 ± 6.02	

Appendix F Table 1. Summary of Adaptive Behavior Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Smith et al, 2000[4] US **IG:** UCLA/Lovaas-based intervention, 15/15 **CG:** Parent training from Lovaas manual, 13/13 **Quality:** Fair	**IG:** 36.07 ± 6.00 **CG:** 35.77 ± 5.77	VABS Composite	**IG:** 63.4 ± 9.35 **CG:** 58.5 ± 16.58	**IG:** 61.19 ± 29.72 **CG:** 58.50 ± 16.58	No statistically significant group difference in any of the VABS measures both at intake and after treatment
		Communication	**IG:** 58.20 ± 5.56 **CG:** 62.00 ± 6.11	**IG:** 67.87 ± 30.08 **CG:** 60.77 ± 17.26	
		Socialization	**IG:** 62.4 ± 7.82 **CG:** 69.15 ± 8.75	**IG:** 66.33 ± 24.78 **CG:** 68.92 ± 16.94	
		Daily living Skills	**IG:** 69.93 ± 8.37 **CG:** 70.62 ± 11.50	**IG:** 62.33 ± 25.76 **CG:** 63.00 ± 16.97	
Non-randomized trials					
Peters-Scheffer et al, 2013[5] Netherlands **IG:** Low intensity Lovaas-based intervention+ specialized preschool, 20/20 **CG:** Eclectic therapy, 20/20 **Quality:** Good	**IG+CG:** 62.52 ± 16.96	Adaptive behavior Composite	**IG:** 18.35 (3.41) **CG:** 19.82 (4.71)	**IG:** 37.35 (13.05) **CG:** 26.71 (9.84), d=1.74, p<0.001	Both groups made significant gains in total adaptive skills (F [1.40, 44.70]= 59.47; p<0.001), but in the treatment group the gains were significantly larger (F [1.40, 44.70]=13.58; p<0.001).
		Communication	**IG:** 23.94 (7.64) **CG:** 24.35 (9.80)	**IG:** 43.71 (17.68) **CG:** 32.35 (14.56), d=1.41	
		Daily living skills	**IG:** 20.82 (6.12) **CG:** 23.00 (9.26)	**IG:** 39.29 (11.13) **CG:** 29.71 (12.15), d=1.62	
		Socialization	**IG:** 19.76 (3.36) **CG:** 22.88 (5.79)	**IG:** 39.35 (10.58) **CG:** 29.71 (9.99), d=2.61	
Peters-Scheffer et al, 2010[6] Netherlands **IG:** Specialized preschool +UCLA/Lovaas-based intervention, 12/12 **CG:** Eclectic therapy, 22/22 **Quality:** Fair	**IG:** 53.5 ± 5.52 **CG:** 52.95 ± 11.14	VABS Composite	**IG:** 20.83 ± 6.69 **CG:** 19.18 ± 4.14	**IG:** 31.75 ± 10.96 **CG:** 22.05 ± 7.47	At pretreatment, no significant group differences in any domain score. Both groups made gains after treatment but gains were larger in the IG group. Composite: F (1,32)=15.68, p<0.01 Communication: F (1,32)=6.48, p=0.02 Daily Living: F (1,32)=13.17, p<0.01 Socialization: F (1,32)=44.86, p<0.01
		Communication	**IG:** 26.92 ± 12.12 **CG:** 25.00 ± 10.00	**IG:** 39.42 ± 15.39 **CG:** 29.95 ± 13.39	
		Daily living skills	**IG:** 23.83 ± 7.28 **CG:** 20.14 ± 4.68	**IG:** 33.25 ± 9.04 **CG:** 23.23 ± 7.70	
		Socialization	**IG:** 20.75 ± 4.54 **CG:** 24.64 ± 8.18	**IG:** 34.08 ± 8.14 **CG:** 25.14 ± 7.21	
Hayward et al, 2009[7, 8] UK **IG:** Intensive clinic-based UCLA/Lovaas-based intervention, 23/20 **CG:** Intensive parent-managed treatment, 21/19 **Quality:** Fair	**IG:** 35.7 ± 6.2 **CG:** 34.4 ± 5.7	VABS Composite	**IG:** 62.3 ± 6.8 **CG:** 65.1 ± 10.4 **IG+CG:** 63.5 ± 8.8	**IG:** 68.4 ± 14.5 **CG:** 72.5 ± 17.3 **IG+CG:** 69.9 ±15.9	No significant group differences in composite measure at intake and at followup. Combining both groups, (IG+CG): Mean change scores: Composite score: 6.4, p<0.001 Communication: 7.6, p<0.01 Daily Living: 1.8, p=NS Socialization: 5.0, p<0.05 Motor: 6.0 , p<0.05

Appendix F Table 1. Summary of Adaptive Behavior Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Reed et al, 2007[9] UK **IG:** High intensity intervention, 14/14 **IGa:** High intensity with focus on Lovaas techniques, 4/4 **IGb:** High intensity with focus on verbal behavior, 5/5 **IGc:** High intensity with focus on CABAS methods, 5/5 **CG:** Low intensity intervention in home-based direct teaching sessions, 13/13 **Quality:** Fair	**IG:** 42.9 (14.8) **IGa:** 47.5 (13.5) **IGb:** 38.0 (9.9) **IGc:** 44.2 (20.5) **CG**: 40.8 (5.6)	VABS composite score	**IG:** 59.3 ± 10.1 **IGa:** 59.8 ± 16.7 **IGb:** 58.2 ± 6.5 **IGc:** 60.0 ± 8.6 **CG:** 56.5 ± 4.4	VABS composite gain score **IG:** NR **IGa:** t<1, ES=0.03 **IGb:** t<1, ES=0.18 **IGc:** t(4)=1.07, ES=0.53 **CG:** NR	Nonsignificant group differences in adaptive behavior, F (2,11)=2.99, p<0.07
Eikeseth et al, 2002[10, 11] Norway **IG:** UCLA/Lovaas-based intervention, 13/13 **CG:** Eclectic therapy, 12/12 **Quality:** Fair	**IG:** 66.31 ± 14.71 **CG:** 65 ± 10.95	VABS Communication	**IG:** 58.23 ± 9.21 **CG:** 63.17 ± 16.11	**IG:** 73.93 ± 16.55 **CG:** 61.58 ± 13.37	No group difference at intake on any of the VABS scores Mean change scores significantly different between groups for communication (p<0.01) and composite scores (p<0.05)
		Daily living skills	**IG:** 56.92 ± 9.8 **CG:** 57.00 ± 15.92	**IG:** 66.15 ± 16.55 **CG:** 62.5 ± 10.97	
		Composite	**IG:** 55.77 ± 8.96 **CG:** 60.00 ± 13.2	**IG:** 67 ± 16.3 **CG:** 60.17 ± 11.69	
		Socialization	**IG:** 59.92 ± 7.19 **CG:** 62.17 ± 10.32	**IG:** 59.92 ± 7.19 **CG:** 70.67 ± 13.66	
Prospective cohort studies					
Eldevik et al. 2012[12] Norway **IG:** Preschool-based early intensive intervention, 31/31 **CG:** Eclectic therapy, 12/12 **Quality:** Fair	**IG:** 42.2 ± 9.0 **CG:** 46.2 ± 12.4	VABS Adaptive behavior composite	**IG:** 62.5 ± 8.2 (46-77) **CG:** 58.9 ± 7.8 (50-73)	**IG:** 68.4 ± 12.6 (46-97) **CG:** 59.6 ± 11.8 (47-83)	CG made significantly larger gains on Composite scores, F (1,39)=4.74, p=0.036, Communication, F (1,38)= 4.82, p=0.034 and socialization, F (1,38) =7.79, p<0.008 Daily living skills were not different between groups, F (1,38)=2.91, p=0.094 Standardized mean difference effect size for change in compact score=0.73 (95% CI, 0.05 to 1.36)
		Communication	**IG:** 61.9 ± 10.2 (48-89) **CG:** 60.0 ± 9.6 (49-81)	**IG:** 70.5 ± 16.9 (42-114) **CG:** 60.0 ± 14.5 (42-84)	
		Daily Living	**IG:** 69.9 ± 10.8 (48-89) **CG:** 64.8 ± 10.6 (54-91)	**IG:** 72.0 ± 12.9 (47-93) **CG:** 63.2 ± 14.2 (48-95)	
		Socialization	**IG:** 63.3 ± 9.8 (49-97) **CG:** 63.1 ± 8.9 (53-82)	**IG:** 69.1 ± 12.0 (49-90) **CG:** 60.8 ± 8.6 (41-80)	

Appendix F Table 1. Summary of Adaptive Behavior Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Eikeseth et al, 2012[13] Norway/Sweden **IG:** Early intensive intervention, 35/13-15 depending on outcome **CG:** Eclectic therapy, 24/NR **Quality:** Fair	**IG:** 3.9 ± 0.9 years **CG:** 4.4 ± 1.2 years	VABS-Total	**IG:** 67.0 ± 10.3 **CG:** 63.6 ± 8.1	**IG:** 75.3 ± 12.0 **CG:** 64.0 ± 12.5	No baseline group differences. At 1 year followup, IG scored significantly higher on all VABS scales compared to CG (p<0.05). Effect sizes for VABS : Total (composite): 0.92 Communication: 1.08 ADL: 0.71 Socialization: 0.75 Motor: 0.70
		Communication	**IG:** 67.1 ± 14.0 **CG:** 65.5 ± 14.2	**IG:** 81.3 ± 16.9 **CG:** 63.6 ± 16.0	
		Activities of daily living	**IG:** 71.8 ± 12.8 **CG:** 67.5 ± 10.9	**IG:** 78.3 ± 14.4 **CG:** 68.0 ± 14.8	
		Motor Skills	**IG:** 75.9 ± 12.8 **CG:** 72.5 ± 10.6	**IG:** 80.6 ± 10.6 **CG:** 71.8 ± 14.4	
Itzchak et al. 2011[14, 15] Israel **IG:** ABA-based approach, 45/45 **CG:** Eclectic approach, 33/33 **Quality:** Fair	**IG:** 25.1 ± 3.9 **CG:** 26.0 ± 4.6	VABS-Socialization standard score	**IG:** 67.8 ± 7.7 **CG:** 70.7 ±7.7	**IG:** 69.6 ± 12.4 **CG:** 77.4 ± 14.4	No significant difference between the 2 groups in VABS scores. IG+CG: Significant predictors of VABS scores at 1 year followup were baseline MSEL-Verbal (p<0001), maternal age (p<0.01), and autism severity x baseline MSEL-Verbal (p<0.05).
		VABS-Communication standard score	**IG:** 67.0 ± 7.8 **CG:** 69.5 ± 10.7	**IG:** 72.9 ± 14.7 **CG:** 78.8 ± 16.2	
		VABS Composite score	**IG:** 66.2 ± 9.6 (49-75) **CG:** 68.6 ± 6.3 (59-81) **IG+CG**: 67.4 ± 6.4	**IG:** NR **CG:** NR **IG+CG**: 68.9 ± 13.0	
		VABS-Daily Living standard score	**IG:** 67.7 ± 7.0 **CG:** 69.4 ± 6.0	**IG:** 67.8 ± 10.9 **CG:** 73.0 ± 14.6	
		VABS-Motor skills standard score	**IG:** 86.2 ± 11.4 **CG:** 88.1 ± 11.0	**IG:** 72.0 ± 12.9 **CG:** 84.5 ± 13.0	
Cohen et al, 2006[16] US **IG:** UCLA/Lovaas-based intervention, 21/21 **CG:** Eclectic therapy, 21/21 **Quality:** Fair	**IG:** 30.2 ± 5.8 **CG:** 33.2 ± 3.7	VABS Communication	**IG:** 69.4 ± 11.8 **CG:** 65.0 ± 6.8	**IG:** NR **CG:** NR **IG/CG:** *P*<0.05	Groups significantly differed in VABS composite score (p<0.01), as well as in individual scores, Communication and Daily Living Skills.
		Composite	**IG:** 69.8 ± 8.1 **CG:** 70.6 ± 9.6	Mean change: **IG:** 9 (n=20) **CG:** -4 (n=20) **IG/CG:** *P*<0.01 Children in the average range, n: **IG:** 8 **CG:** 3 **IG/CG:** *P*=0.10	
		Daily Living skills	**IG:** 73.2 ± 9.2 **CG:** 72.7 ± 12.5	**IG:** NR **CG:** NR **IG/CG:** *P*<0.05	
		Social skills	**IG:** 70.3 ± 10.9 **CG:** 75.1 ± 13.0	**IG:** NR **CG:** NR **IG/CG:** *P*<0.10	

Appendix F Table 1. Summary of Adaptive Behavior Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Followup Score, Mean±SD	Analytic Data
Howard et al. 2005[17] US **IG:** UCLA/Lovaas-based intervention, 37/29 **CG1:** Intensive eclectic therapy **CG2:** Non-intensive eclectic therapy **CG2+CG3:** 41/32 **Quality:** Fair	**IG:** 30.86 ± 5.16 **CG1:** 37.44 ± 5.68 **CG2:** 34.56 ± 6.53	VABS-Composite	**IG:** 70.46 ± 11.85 **CG1:** 69.81 ± 10.48 **CG2:** 71.62 ± 10.47	**IG:** 81.32 ± 11.14 **CG1:** 69.25 ± 12.91 **CG2:** 68.25 ± 9.86	Significant difference in composite scores between IG and CG1/CG2, p<0.01 Significant difference in Self-help scores between IG and CG1/CG2, p<0.01
		Denver Developmental Screening Test II, DP-II, RIDES for Self-help	**IG:** 70.71 ± 10.14 **CG1:** 68.06 ± 11.61 **CG2:** 73.43 ± 10.39	**IG:** 76.56 ± 11.59 **CG1:** 70.00 ± 11.92 **CG2:** 65.19 ± 8.84	
Retrospective cohort studies					
Flanagan et al, 2012[18, 19] Canada **IG:** Intensive behavioral intervention, 61/61 **CG:** Eclectic therapy 61/61 **Quality:** Fair	**IG:** 42.93 ± 11.53 **CG:** 42.79 ± 10.51	VABS standard scores composite	**IG:** 55.38 ± 7.00 **CG:** 55.49 ± 7.11	**IG:** 56.34 ± 14.40 **CG:** 52.19 ± 8.77	Estimated marginal scores: **IG:** 56.96 **CG:** 50.66 (p=0.008), d=0.53
		Ratio scores Composite	**IG:** 30.78 ± 10.78 **CG:** 30.79 ± 10.67	**IG:** 41.77 ± 20.26 **CG:** 31.15 ± 11.82	**IG:** 40.75 **CG:** 30.32 (p=0.002), d=0.63
		Communication	**IG:** 25.47 ± 15.81 **CG:** 25.50 ± 11.97	**IG:** 46.60 ± 29.91 **CG:** 30.33 ± 16.98	**IG:** 43.45 **CG:** 29.80 (p=0.006), d=0.56
		Daily Living Skills	**IG:** 42.79 ± 11.97 **CG:** 42.87 ± 12.11	**IG:** 44.83 ± 14.01 **CG:** 40.03 ± 11.06	**IG:** 45.04 **CG:** 38.80 (p=0.023), d=0.49
		Socialization	**IG:** 24.08 ± 9.36 **CG:** 23.99 ± 11.22	**IG:** 33.90 ± 19.04 **CG:** 23.11 ± 10.85	**IG:** 33.49 **CG:** 21.88 (p=0.001), d=0.75

Abbreviations: ABA=Applied Behavior Analysis; ADL=Activities of Daily Living; CABAS=Comprehensive Application of Behavior Analyst; CG=Control Group; DP-II=Developmental Profile II; ESDM=Early Start Denver Model; IG=Intervention Group; MSEL=Mullen Scales of Early Learning; NR=Not Reported; VABS=Vineland Adaptive Behavior Scale.

Appendix F Table 2. Summary of Symptom Severity Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
RCTs					
Strain et al, 2011[20] US **IG:** LEAP program with coaching and training, 28 classrooms (27 analyzed)/ 177 children **CG:** LEAP intervention manuals only, 28 classrooms (23 analyzed)/ 117 children **Quality:** Fair	**IG:** 50.1 ± 4.6 **CG:** 50.7 ± 4.2	CARS	**IG:** 39.0 ± 6.2 **CG:** 37.4 ± 5.9	**IG:** 32.9 ± 3.9 **CG:** 34.6 ± 4.2	Children in IG showed an average reduction in severity of 6.1 points as compared with an average reduction of 2.8 points for CG class children, p<0.05; ES=0.59
Dawson et al, 2012[1, 2] US **IG:** ESDM, 24/24 **CG:** Eclectic therapy, 24/21 Quality: Good	**IG:** 23.9 ± 4.0 **CG:** 23.1 ± 3.9	ADOS	**G1:** 7.2 ± 1.7 **G2:** 6.9 ± 1.7	**G1:** 7.0 ± 1.9 **G2:** 7.3 ± 1.8	No group difference in ADOS severity scores. Group x Time (baseline vs. 2 year): F=3.29, p=0.422
Nonrandomized trials					
Peters-Scheffer et al, 2013[5] Netherlands **IG:** Low intensity Lovaas-based intervention+ specialized preschool, 20/20 **CG:** Eclectic therapy, 20/20 **Quality:** Good	**IG+CG:** 62.52 ± 16.96 (median)	ADOS-total score	**IG:** 17.00 ± 3.28 **CG:** 15.45 ± 2.72	**IG:** 12.05 ± 5.41 **CG:** 15.15 ± 4.26	Fewer autistic symptoms observed in IG than CG at followup Effect size: autism severity ADOS: Cohen's d=1.51 CARS: Cohen's d=1.50
		CARS-total score	**IG:** 43.84 ± 4.30 **CG:** 40.79 ± 6.20	**IG:** 34.89 ± 3.62 **CG:** 39.95 ± 4.62	
Peters-Scheffer et al, 2010[6] Netherlands **IG:** Specialized preschool +UCLA/Lovaas-based intervention, 12/12 **CG:** Eclectic therapy, 22/22 **Quality:** Fair	**IG:** 53.5 ± 5.52 **CG:** 52.95 ± 11.14	PDD-MRS	**IG:** 11.58 ± 4.42 **CG:** 12.91 ± 3.79	**IG:** 10.25 ± 3.14 **CG:** 11.27 ± 3.84	No significant group differences in symptom severity at pretreatment, t(20)=0.88, p=0.39, and posttreatment, t(27)=0.84, p=0.41. Decrease of symptom severity over time in both groups, F(1,32)=6.22, p=0.02

Appendix F Table 2. Summary of Symptom Severity Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
Reed et al, 2007[9] UK **IG:** High intensity intervention, 14/14 **IGa:** High intensity with focus on Lovaas techniques, 4/4 **IGb:** High intensity with focus on verbal behavior, 5/5 **IGc:** High intensity with focus on CABAS methods, 5/5 **CG:** Low intensity intervention in home-based direct teaching sessions, 13/13 **Quality:** Fair	**IG:** 42.9 (14.8) **IGa:** 47.5 (13.5) **IGb:** 38.0 (9.9) **IGc:** 44.2 (20.5) **CG**: 40.8 (5.6)	GARS Autism Quotient	**IG:** 89.1 ± 14.7 **IGa:** 93.0 ± 19.9 **IGb:** 87.6 ± 11.1 **IGc:** 87.4 ± 16.1 **CG:** 95.1 ± 11.6	Mean change ± SD: **IG:** -2.2 ± 7.8 **CG:** 1.6 ± 6.2	No significant difference between the groups, t(25)=1.41, p>0.10
Prospective cohort studies					
Boyd et al, 2013[21] US **IG1:** TEACCH preschools, 85/81 **IG2:** LEAP preschools, 54/48 **CG:** Nonmodel specific preschools, 59/56 **Quality:** Fair	**IG1:** 48 ± 6.84 **IG2:** 47.52 ± 8.4 **CG:** 48.84 ± 7.68	Autism characteristics and severity (ACS)	**IG1:** -0.11 ± 0.76 **IG2:** 0.066 ± 0.765 **CG:** 0.381 ± 0.859	**IG1:** -0.299 ± 0.928 **IG2:** -0.144 ± 0.837 **CG:** 0.124 ± 0.866	Significant baseline group differences (p=0.0013) All groups showed significant change from baseline (p<0.05), but there is no between group differences in severity
Zachor et al, 2007[22] Israel **IG:** UCLA/Lovaas-based intervention, 53/53 **CG:** Eclectic therapy, 15/15 **Quality:** Fair	**IG:** 25.1 ± 3.8 **CG:** 26.3 ± 4.6	ADOS-Language & communication	**IG:** 13.8 ± 4.3 **CG:** 11.8 ± 4.3	**IG:** 7.2 ± 4.1 **CG:** 9.7 ± 3.0	No significant difference between groups at preintervention time in ADOS scores (F(2,36)=1.05, p=0.359, η^2=0.055). Significant group differences for language & communication subscale, F(2,38)=9.59, p<0.01, η^2=0.206. No significant group differences for reciprocal social interaction, F(2,38)=3.39, η^2=0.074.
		Reciprocal social interaction	**IG:** 17.9 ± 6.2 **CG:** 16.3 ± 5.2	**IG:** 11.1 ± 6.7 **CG:** 13.3 ± 4.8	

Appendix F Table 2. Summary of Symptom Severity Outcomes in Direct to Child Provision Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
Howard et al, 2005[17] US **IG:** UCLA/Lovaas-based intervention, 37/29 **CG1:** Intensive eclectic therapy **CG2:** Nonintensive eclectic therapy **CG2+CG3:** 41/32 **Quality:** Fair	**IG:** 30.86 ± 5.16 **CG1:** 37.44 ± 5.68 **CG2:** 34.56 ± 6.53	Number of DSM-IV criteria	**IG:** 7.55 ± 1.39 **CG1:** 7.27 ± 1.56 **CG2:** 7.33 ± 2.02	NR	IG vs CG1/CG2: mean difference = 0.25, p=NS CG1 vs. CG2 , mean difference = -0.06, p=NS
Retrospective cohort studies					
Flanagan et al, 2012[18, 19] Canada **IG:** Intensive behavioral intervention, 61/61 **CG:** Eclectic therapy 61/61 **Quality:** Fair	**IG:** 42.93 ± 11.53 **CG:** 42.79 ± 10.51	CARS	**IG:** 32.83 ± 3.99 **CG:** 32.62 ± 3.74	**IG:** 30.20 ± 4.97 **CG:** 32.57 ± 5.55	No baseline group difference in CARS (total score), t=-0.29, p=0.77 Significant group difference at time 2: F=4.64, p=0.033, d=0.53

Abbreviations: ACS=Autism Characteristics and Severity; ADOS=Autism Diagnostic Observation Schedule; CARS=Childhood Autism Rating Scale; CG=Control Group; ESDM=Early Start Denver Model; GARS=Gilliam Autism Rating Scale; IG=Intervention Group; LEAP=Learning Experiences and Alternate Program for preschoolers and their parents; SD=Standard Deviation.

Appendix F Table 3. Summary of Adaptive Behavior Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
RCTs					
Schreibman et al, 2013[23] US **IG:** Pivotal Response Training, 20/20 **CG:** PECS, 19/19 **Quality:** Good	**IG:** 29.5 ± 6.9 **CG:** 28.9 ± 4.2	VABS, Communication	**IG:** 62.2 ± 4.7 **CG:** 60.2 ± 7.5	Followup: **IG:** 68.4 ± 14.5 **CG:** 62.6 ± 12.7	Change over time and across conditions: Time effect: F=4.09, p=0.037, ES=0.11 Rx effect: F=2.263, p=0.142 Time x Rx: F=1.765, p=0.19
Landa et al, 2012[24, 25] US **IG:** Assessment Evaluation and Programming System for Infants and Children (AEPS) curriculum + additional joint attention and social interaction opportunities, 25/24 **CG:** AEPS curriculum, 25/24 **Quality:** Good	**IG:** 28.6 ± 2.6 **CG:** 28.8 ± 2.8	VABS, Communication standard score	**IG+CG:** n=46, 69.7 ± 9.6	**IG+CG:** n=48, 82.4 ± 20.4, d=0.38	Significant change score (IG+CG): n=46, 12.7 ± 19.4, d=0.81, p<0.001
Roberts et al, 2011[26] Australia **IG:** Individualized home-based program, 34/27 **CG1:** Small group center-based program combined with parent training and support group, 33/29 **CG2:** Waitlist, 28/28 **Quality:** Good	**IG:** 41.5 **CG1:** 43.1 **CG2:** 43.7	VABS-Social	**IG:** 68.7 ± 7.3 **CG1:** 70.1 ± 7.3 **CG2:** 70.8 ± 9.9	**IG:** 66.4 ± 7.7 **CG1:** 72.6 ± 11.2 **CG2:** 73.1 ± 10.8	Significant pair-wise group differences on VABS social scale, Mean difference (95% CI): IG vs. CG1: 5.2 (0.7 to 9.6), p=0.02 IG vs. CG2: 5.2 (-0.7 to 9.7), p=0.02 CG1 vs. CG2: 0.1 (-4.3 to 4.4), p=0.98 3-group comparison: p=0.03 No significant communication score difference between groups (p>0.05)
		VABS communication	**IG:** 64.4 ± 12.8 **CG1:** 66.9 ± 12.5 **CG2:** 68.5 ± 17.0	**IG:** 68.4 ± 15.6 **CG1:** 76.1 ± 17.1 **CG2:** 74.2 ± 15.5	
Aldred et al, 2012[27, 28] UK **IG:** Parent training in social communication intervention plus community intervention 14/14 **CG:** Eclectic therapy, 14/14 **Quality:** Good	**IG:** 51.4 ± 11.8 **CG:** 50.9 ± 16.3	VABS communication	**IG:** 22.6 ± 13.3 **CG:** 20.0 ± 10.8	**IG:** 36.9 ± 21.2 **CG:** 28.7 ± 16.6	Nonsignificant mean communication change score (IG: 14.3, CG: 8.7) after adjusting for baseline score (F[1,25]=2.58, p=0.121)

Appendix F Table 3. Summary of Adaptive Behavior Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
Rogers et al, 2012[29, 30] US **IG:** Parent-delivered Early Start Denver model (ESDM) 49/49 **CG:** Eclectic therapy, 49/49 **Quality:** Fair	**IG:** 21.02 ± 3.51 **CG:** 20.94 ± 3.42	VABS II, Communication	**IG:** 67.66 ± 13.19 **CG:** 67.29 ± 11.05	**IG:** 72.55 ± 12.06 **CG:** 74.29 ± 14.55	No group differences in any of the adaptive measures after 12 week intervention Group effect size (d): Communication: IG: 0.69; G2: 0.84 Daily living: IG: -0.08; G2: 0.08 Socialization: IG: 0.08, G2: 0.07 Adaptive behavior composite: IG: 0.1, G2: 0.29
		VABS Daily Living Skills	**IG:** 83.07 ± 12.4 **CG:** 83.21 ± 10.6	**IG:** 82.25 ± 13.82 **CG:** 84.04 ± 13.5	
		VABS Socialization	**IG:** 76.68 ± 8.74 **CG:** 77.95 ± 8.01	**IG:** 77.32 ± 9.19 **CG:** 78.67 ± 10.78	
		VABS Adaptive Behavior Composite	**IG:** 76.76 ± 10.3 **CG:** 78.22 ± 8.88	**IG:** 77.43 ± 9.59 **CG:** 80.33 ± 11.34	
Green et al, 2010[31] UK **IG:** Preschool autism communication intervention (PACT), 77/74 **CG:** Eclectic therapy, 75/72 **Quality:** Fair	**IG:** 45 **CG:** 45	VABS Communication	NR	**IG:** 64.3 ± 17.7 **CG:** 67.7 ± 17.5	No significant group differences with the teacher Vineland Communication (difference of -3·52 [-7·55 to 0·52]) and Adaptive Behavior Composite standard scores (difference of -2·76 [-6·65 to 1·14])
		Adaptive behavior composite	NR	**IG:** 60.3 ± 15.2 **CG:** 62.8 ± 14.8	
Prospective cohort studies					
Strauss et al, 2012[32, 33] Italy **IG:** Staff & parent mediated early intervention, 24/24 **CG:** Eclectic therapy, 20/20 **Quality:** Good	**IG:** 55.67 ± 17.63 **CG:** 41.94 ± 13.07	VABS-Standard Scores socialization	**IG:** 61.96 ± 21.31 **CG:** 56.88 ± 19.21	**IG:** 67.78 ± 19.93 **CG:** 70.50 ± 24.04	CG achieved greater gains in adaptive behavior socialization, t(19)=3.434, p≤0.01. IG did not show an increase in socialization skills after 6 months. IG & CG: Significant change from baseline communication & daily living scores (p<0.004). IG: No significant change from baseline for motor skills (p=0.079) but CG had significant change from baseline (p=0.001). IG & CG: Significant change from baseline ABC scores (p≤0.001). No significant group differences in adaptive measures.
		Communication	**IG:** 71.00 ± 39.24 **CG:** 60.78 ± 30.42	**IG:** 91.43 ± 40.44 **CG:** 83.56 ± 41.32	
		Daily living	**IG:** 78.43 ± 33.39 **CG:** 56.44 ± 23.81	**IG:** 100.26 ± 35.60 **CG:** 88.33 ± 37.29	
		Motor skills	**IG:** 105.78 ± 22.38 **CG:** 92.00 ± 19.97	**IG:** 112.87 ± 13.30 **CG:** 106.59 ± 21.63	
		AB-Composite	**IG:** 79.29 ± 22.84 **CG:** 66.92 ± 19.25	**IG:** 93.09 ± 23.61 **CG:** 84.88 ± 29.03	
Reed et al, 2012[34] UK **IG:** ABA, 14 **CG1:** Special nursery, 21 **CG2:** Portage, 18 **CG3:** Local authority-developed parent training, 13 **Quality:** Fair	**IG:** 39.0 ± 6.9 **CG1:** 41.5 ± 4.0 **CG2:** 39.5 ± 6.3 **CG3:** 40.2 ± 6.3	VABS composite	**IG:** 58.4 ± 10.6 **CG1:** 53.3 ± 4.2 **CG2:** 56.6 ± 7.0 **CG3:** 54.0 ± 4.5	VABS composite change score **IG:** 11.9 ± 7.7 **CG1:** 6.8 ± 15.7 **CG2:** 2.5 ± 6.1 **CG3:** 2.7 ± 8.7	Significant difference between the standardized beta coefficient for the IG and CG2, p<0.05. None of the other standardized beta coefficients differed between the groups, p values >0.1.

Appendix F Table 3. Summary of Adaptive Behavior Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months±SD	Measure	Baseline Score, Mean±SD	Follow-Up Score, Mean±SD	Analytic Data
Keen et al, 2010[35] Australia **IG:** Professional parent intervention, 17 families/NR **CG:** Self-directed video based parent intervention, 22 families/NR **Quality:** Good	**IG:** 36.38 ± 7.54 **CG:** 35.71 ± 6.92	SIB-R	**IG:** 451.6 ± 8.12 **CG:** 448.1 ± 9.25	**IG:** 457 ± 7.62 **CG:** 452.1 ± 11.02	No significant group difference after controlling for baseline measures (Baseline x IG: beta=-0.35, SE=0.179, 95% CI: -0.70 to 0.0008)

Abbreviations: ABC=Autism Behavior Checklist; AEPS=Assessment Evaluation and Programming System for infants and children; CARS=Childhood Autism Rating Scale; CI=Confidence Interval; ES=effect size; LEAP=Learning Experiences and Alternate Program for preschoolers and their parents; NR=Not Reported; PACTS=Preschool Autism Communication Intervention; PDD-NOS=Pervasive Developmental Disorder-Not Otherwise Specified; PECS=Picture Exchange Communication System; SIB-R=Scales of Independent Behavior; TEACCH=Treatment and Education of Autistic and related Communication Handicapped Children; VABS=Vineland Adaptive Behavior Scale.

Appendix F Table 4. Summary of Symptom Severity Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months ± SD	Measure	Baseline Score, Mean ± SD	Follow-Up Score, Mean ± SD	Analytic Data
RCTs					
Aldred et al, 2012[27, 28] UK **IG:** Parent training in social communication intervention plus community intervention, 14/14 **CG:** Eclectic therapy, 14/14 **CGa:** aged 24-27 months, total ADOS score 11-17 (young high functioning) **CGb:** aged 24-47 months; total ADOS score 18-24 (young low functioning) **CGc:** aged 48-71 months; total ADOS score 11-17 (older high functioning) **CGd:** aged 48-71 months; total ADOS score 18-24 (older low functioning) **Quality:** Good	**IG:** 51.4 ± 11.8 **CG:** 50.9 ± 16.3	ADOS total score	**IG:** 16.1 ± 4.5 **IGa:** 12 ± 3.3 **IGb:** 19 ± 1.3 **IGc:** 14 ± 3.3 **IGd:** 20 ± 1 **CG:** 15.6 ± 4.9 **CGa:** 11 ± 2.3 **CGb:** 19 ± 1 **CGc:** 14 ± 3.3 **CGd:** 20 ± 1.3	**IG:** 11.8 ± 6.4 **IGa:** 6 ± 3.6 **IGb:** 13 ± 5.6 **IGc:** 11 ± 4.5 **IGd:** 17 ± 2.6 **CG:** 16.1 ± 4.4 **CGa:** 13 ± 4 **CGb:** 16 ± 4.3 **CGc:** 16 ± 1.3 **CGd:** 20 ± 0.6	Co-varying for baseline ADOS score, there was a significant difference in ADOS change between the groups (F[1,25]=7.30; p=0.01).
Oosterling et al, 2010[36] **IG:** Nonintensive parent training+specialized preschool, 40/36 **CG:** Eclectic therapy, 35/31 **Quality:** Fair	**IG:** 35.2 ± 5.5 **CG:** 33.3 ± 6.4	ADOS Joint attention factor	NR	Change from baseline scores: **G1:** -0.8 ± 2.3 **G2:** -0.9 ± 0.2	Joint attention factor: Group effect, F=0.12 Time effect, F=0.67, Group*Time effect= 0.76
		Social affect	NR	**G1:** -2.5 ± 4.0 **G2:** -2.3 ± 3.7	Social affect: Group effect, F=0.01 Time effect, F=6.08 (p<0.05) Group* Time effect, F= 0.10
Rogers et al, 2012[29, 30] US **IG:** Parent-delivered Early Start Denver model (ESDM), 49/49 **CG:** Eclectic therapy, 49/49 **Quality:** Fair	**IG:** 21.02 ± 3.51 **CG:** 20.94 ± 3.42	Modified ADOS social affect	**IG:** 29.45 ± 9.16 **CG:** 34.14 ± 8.69	**IG:** 26.61 ± 10.14 **CG:** 27.33 ± 10.62	Change from baseline IG: d=- 0.37 CG: d=- 0.63 CG showed greater improvement in social affect score than IG (estimated difference between groups =3.43, SD=1.72, p=0.05).
		ADOS–Restrictive and Repetitive	**IG:** 3.92 ± 2.01 **CG:** 4.31 ± 1.92	**IG:** 3.96 ± 1.86 **CG:** 3.82 ± 2.04	Change from baseline: IG: d=0.02 CG: d=-0.22 No significant group difference in ADOS scores No effect of intervention hours on the group difference in both scores, p>0.05

Appendix F Table 4. Summary of Symptom Severity Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months ± SD	Measure	Baseline Score, Mean ± SD	Follow-Up Score, Mean ± SD	Analytic Data
Pajareya et al, 2011[37] Thailand **IG:** DIR/Floortime,16/15 **CG:** Usual care, 16/16 **Quality:** Fair	**IG:** 56.6 ± 10.1 **CG:** 51.5 ± 13.9	CARS	**IG:** 37.2 ± 6.2 **CG:** 39.7 ± 6.6	Change scores: **IG:** 2.9 ± 2.0 **CG:** 0.8 ±1.2	No baseline difference between the groups, p=0.86. Change CARS scores: significantly greater decrease for the IG as compared to CG (p=0.002)
Carter et al, 2011[38] US **IG:** More than Words, 32/29 **CG:** Eclectic therapy, 30/26 **Quality:** Fair	**IG:** 21.11 ± 2.71 **CG**: 21.51 ± 2.82	PIA-CV nonverbal communication raw scores	**IG:** 2.30 ± 0.64 **CG:** 2.28 ± 0.73	**IG:** 2.89 ± 0.67 **CG:** 2.92 ± 0.65	Residualized gain scores from baseline: **IG: -**0.05 ± 0.63 **CG:** 0.06 ± 0.58 ES=-0.19 , 95%CI :-0.81 to 0.43
Green et al. 2010[31] UK **IG:** Preschool autism communication intervention (PACT), 77/74 **CG:** Eclectic therapy, 75/72 Quality: Fair	**IG:** 45 **CG:** 45	ADOS-G Total social communication algorithm score	**IG:** 19.6 ± 4.2 **CG:** 19.3 ± 4.0	**IG:** 15.7 ± 6.0 **CG:** 16.5 ± 5.7 Change from baseline, mean (SD): **IG:** -3.9 ± 4.7 **CG:** -2.9 ± 3.9	Change in ADOS-G diagnosis to autism spectrum disorder, n (%): IG: 22 (30) CG: 17 (24) Change in ADOS-G diagnosis to non-spectrum, n (%): IG: 4 (5) CG: 5 (7) Treatment effect with ANCOVA estimates: without baseline covariate adjustment: -1·06 (95% CI, -2·48 to 0·36) with adjustment for center and age: -1·00 (-2·38 to 0·39) No interaction with treatment for ADOS-G algorithm score >17 (p=0.85)
Drew et al, 2002[39] UK **IG:** Parent training, 12/12 **CG:** Local/eclectic services 12/12 **Quality:** Fair	**IG:** 21.4 ± 2.7 **CG:** 23.6 ± 3.8	ADI-R Non-Verbal Communication	**IG:** 12.8 ± 1.6 **CG:** 12 ± 2.4	**IG:** 11.0 ± 2.8 **CG:** 11.9 ± 1.8	No significant group differences in any of the Severity measures (p>0.05)
		Reciprocal Social Interaction	**IG:** 19.6 ± 3.0 **CG:** 20.3 ± 4.5	**IG:** 18.3 ± 4.9 **CG:** 20.1 ± 4.3	
		Repetitive & Stereotyped Behavior	**IG:** 3.2 ± 1.1 **CG:** 3.7 ± 1.6	**IG:** 3.9 ± 1.8 **CG:** 4.2 ± 2.0	
Prospective cohort studies					

Appendix F Table 4. Summary of Symptom Severity Outcomes in Parent Training Studies

Author, Year, Country Groups, N Enrollment/ N Final	Age at Intake Mean months ± SD	Measure	Baseline Score, Mean ± SD	Follow-Up Score, Mean ± SD	Analytic Data
Strauss et al, 2012[32, 33] Italy **IG:** Staff & parent mediated early intervention, 24/24 **CG:** Eclectic therapy, 20/20 **Quality:** Good	**IG:** 55.67 ± 17.63 **CG:** 41.94 ± 13.07	ADOS-Total	**IG:** 15.96 ± 4.33 **CG:** 14.56 ± 5.05	**IG:** 13.21 ± 3.83 **CG:** 13.56 ± 4.72	Change from baseline: IG: t=-3.1, p=0.005 CG: t=-1.826, p=0.09
		ADOS Social interaction	**G1:** 10.54 ± 2.34 **G2:** 9.63 ± 3.24	**G1:** 8.83 ± 2.70 **G2:** 9.00 ± 2.97	IG: t= - 3.995, p<0.001 CG: t= -1.775, p=0.096
		ADOS Communication	**G1:** 6.04 ± 1.88 **G2:** 4.94 ± 2.23	**G1:** 4.38 ± 1.34 **G2:** 4.56 ± 1.97	IG: t=- 3.745, p<0.001 CG: t=- 1.031, p=0.319 No age effect on score change differences on ADOS total, F(44)=1.009, p=0.230. In both groups, the predictive power of parental stress on autism severity was modified by perception of difficult child, with higher perceptions of difficulty associated with lower decreases in autism severity.

Abbreviations: ABA=Applied Behavior Analysis; ADI-R=Autism Diagnostic Interview Revised; ADOS=Autism Diagnostic Observation Schedule; ADOS-G=Autism Diagnostic Observation Schedule-Generic; CARS=Childhood Autism Rating Schedule; CG=Control Group; ESDM=Early Start Denver Model; ES=effect size; GARS=Gilliam Autism Rating Scale; IG=Intervention Group; NR=Not Reported; PACT=Preschool Autism Communication Intervention; PDD-NOS=Pervasive Developmental Disorder-Not Otherwise Specified; PIA-CV=Parent Interview for Autism-Clinical Version; SD=Standard Deviation.

References

1. Dawson G, Jones EJ, Merkle K, et al. Early behavioral intervention is associated with normalized brain activity in young children with autism. J Am Acad Child Adolesc Psychiatry 2012 Nov;51(11):1150-9. PMID: 23101741.
2. Dawson G, Rogers S, Munson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics 2010 Jan;125(1):e17-23. PMID: 19948568.
3. Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four-year outcome and predictors. Am J Ment Retard 2005 Nov;110(6):417-38. PMID: 16212446.
4. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. Am J Ment Retard 2000 Jul;105(4):269-85. PMID: 10934569.
5. Peters-Scheffer N, Didden R, Mulders M, et al. Effectiveness of low intensity behavioral treatment for children with autism spectrum disorder and intellectual disability. Research in Autism Spectrum Disorders 2013;7(9):1012-25.
6. Peters-Scheffer N, Didden R, Mulders M, et al. Low intensity behavioral treatment supplementing preschool services for young children with autism spectrum disorders and severe to mild intellectual disability. Res Dev Disabil 2010 Nov-Dec;31(6):1678-84. PMID: 20627451.
7. Hayward D, Eikeseth S, Gale C, et al. Assessing progress during treatment for young children with autism receiving intensive behavioural interventions. Autism 2009 Nov;13(6):613-33. PMID: 19933766.
8. Eikeseth S, Hayward D, Gale C, et al. Intensity of Supervision and Outcome for Preschool Aged Children Receiving Early and Intensive Behavioral Interventions: A Preliminary Study. Research in Autism Spectrum Disorders 2009 Jan;3(1):67-73.
9. Reed P, Osborne LA, Corness M. Brief report: relative effectiveness of different home-based behavioral approaches to early teaching intervention. J Autism Dev Disord 2007 Oct;37(9):1815-21. PMID: 17180714.
10. Eikeseth S, Smith T, Jahr E, et al. Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. Behav Modif 2002 Jan;26(1):49-68. PMID: 11799654.
11. Eikeseth S, Smith T, Jahr E, et al. Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: a comparison controlled study. Behav Modif 2007 May;31(3):264-78. PMID: 17438342.

12. Eldevik S, Hastings RP, Jahr E, et al. Outcomes of behavioral intervention for children with autism in mainstream pre-school settings. J Autism Dev Disord 2012 Feb;42(2):210-20. PMID: 21472360.
13. Eikeseth S, Klintwall L, Jahr E, et al. Outcome for Children with Autism Receiving Early and Intensive Behavioral Intervention in Mainstream Preschool and Kindergarten Settings. Research in Autism Spectrum Disorders 2012;6(2):829-35.
14. Itzchak EB, Zachor DA. Who Benefits from Early Intervention in Autism Spectrum Disorders? Research in Autism Spectrum Disorders 2011;5(1):345-50. P
15. Zachor DA, Itzchak EB. Treatment Approach, Autism Severity and Intervention Outcomes in Young Children. Research in Autism Spectrum Disorders 2010;4(3):425-32.
16. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S145-55. PMID: 16685181.
17. Perry A, Cummings A, Geier JD, et al. Predictors of Outcome for Children Receiving Intensive Behavioral Intervention in a Large, Community-Based Program. Research in Autism Spectrum Disorders 2011;5(1):592-603.
18. Flanagan HE, Perry A, Freeman NL. Effectiveness of large-scale community-based intensive Behavioral Intervention: A waitlist comparison study exploring outcomes and predictors. Research in Autism Spectrum Disorders 2012;6(2):673-82.
19. Perry A, Blacklock K, Dunn Geier J. The relative importance of age and IQ as predictors of outcomes in Intensive Behavioral Intervention. Research in Autism Spectrum Disorders 2013;7(9):1142-50.
20. Strain PS, Bovey EH. Randomized, Controlled Trial of the Leap Model of Early Intervention for Young Children with Autism Spectrum Disorders. Topics in Early Childhood Special Education 2011;31(3):133-54.
21. Boyd BA, Hume K, McBee MT, et al. Comparative Efficacy of LEAP, TEACCH and Non-Model-Specific Special Education Programs for Preschoolers with Autism Spectrum Disorders. J Autism Dev Disord 2013 Jun 28PMID: 23812661.
22. Zachor DA, Ben-Itzchak E, Rabinovich A-L, et al. Change in autism core symptoms with intervention. Research in Autism Spectrum Disorders 2007;1(4):304-17.
23. Schreibman L, Stahmer AC. A Randomized Trial Comparison of the Effects of Verbal and Pictorial Naturalistic Communication Strategies on Spoken Language for Young Children with Autism. J Autism Dev Disord 2013 Nov 23PMID: 24272416.
24. Landa RJ, Holman KC, O'Neill AH, et al. Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: a randomized controlled trial. J Child Psychol Psychiatry 2011 Jan;52(1):13-21. PMID: 21126245.
25. Landa RJ, Kalb LG. Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention. Pediatrics 2012 Nov;130 Suppl 2:S186-90. PMID: 23118250.
26. Roberts J, Williams K, Carter M, et al. A Randomised Controlled Trial of Two Early Intervention Programs for Young Children with Autism: Centre-Based with Parent Program and Home-Based. Research in Autism Spectrum Disorders 2011;5(4):1553-66.
27. Aldred C, Green J, Adams C. A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness. J Child Psychol Psychiatry 2004 Nov;45(8):1420-30. PMID: 15482502.
28. Aldred C, Green J, Emsley R, et al. Mediation of treatment effect in a communication intervention for pre-school children with autism. Journal of Autism and Developmental Disorders 2012;42(3):447-54.
29. Rogers SJ, Estes A, Lord C, et al. Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. Journal of the American Academy of Child & Adolescent Psychiatry 2012;51(10):1052-65.
30. Estes A, Vismara L, Mercado C, et al. The Impact of Parent-Delivered Intervention on Parents of Very Young Children with Autism. J Autism Dev Disord 2013 Jul 10PMID: 23838727.
31. Green J, Charman T, McConachie H, et al. Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. Lancet 2010 May 20PMID: 20494434.
32. Strauss K, Vicari S, Valeri G, et al. Parent inclusion in Early Intensive Behavioral Intervention: the influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. Res Dev Disabil 2012 Mar-Apr;33(2):688-703. PMID: 22188793.
33. Fava L, Strauss K, Valeri G, et al. The Effectiveness of a Cross-Setting Complementary Staff- and Parent-Mediated Early Intensive Behavioral Intervention for Young Children with ASD. Research in Autism Spectrum Disorders 2011;5(4):1479-92.
34. Reed P, Osborne L. Impact of Severity of Autism and Intervention Time-Input on Child Outcomes: Comparison across Several Early Interventions. British Journal of Special Education 2012;39(3):130-6.

35. Keen D, Couzens D, Muspratt S, et al. The effects of a parent-focused intervention for children with a recent diagnosis of autism spectrum disorder on parenting stress and competence. Research in Autism Spectrum Disorders 2010;4(2):229-41.
36. Oosterling I, Visser J, Swinkels S, et al. Randomized controlled trial of the focus parent training for toddlers with autism: 1-year outcome. J Autism Dev Disord 2010 Dec;40(12):1447-58. PMID: 20440639.
37. Pajareya K, Nopmaneejumruslers K. A pilot randomized controlled trial of DIR/Floortime parent training intervention for pre-school children with autistic spectrum disorders. Autism 2011 Sep;15(5):563-77. PMID: 21690083.
38. Carter AS, Messinger DS, Stone WL, et al. A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. Journal of Child Psychology and Psychiatry 2011;52(7):741-52.
39. Drew A, Baird G, Baron-Cohen S, et al. A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. Eur Child Adolesc Psychiatry 2002 Dec;11(6):266-72. PMID: 12541005.