Screening for Speech and Language Delay in Children 5 Years Old and Younger: A Systematic Review

Ina F. Wallace, PhD[a], Nancy D. Berkman, PhD[a], Linda R. Watson, EdD[b], Tamera Coyne-Beasley, MD, MPH[c,d], Charles T. Wood, MD[c], Katherine Cullen, BA[a], Kathleen N. Lohr, PhD[a]

abstract

BACKGROUND AND OBJECTIVES: No recommendation exists for or against routine use of brief, formal screening instruments in primary care to detect speech and language delay in children through 5 years of age. This review aimed to update the evidence on screening and treating children for speech and language since the 2006 US Preventive Services Task Force systematic review.

METHODS: Medline, the Cochrane Library, PsycInfo, Cumulative Index to Nursing and Allied Health Literature, ClinicalTrials.gov, and reference lists. We included studies reporting diagnostic accuracy of screening tools and randomized controlled trials reporting benefits and harms of treatment of speech and language. Two independent reviewers extracted data, checked accuracy, and assigned quality ratings using predefined criteria.

RESULTS: We found no evidence for the impact of screening on speech and language outcomes. In 23 studies evaluating the accuracy of screening tools, sensitivity ranged between 50% and 94%, and specificity ranged between 45% and 96%. Twelve treatment studies improved various outcomes in language, articulation, and stuttering; little evidence emerged for interventions improving other outcomes or for adverse effects of treatment. Risk factors associated with speech and language delay were male gender, family history, and low parental education. A limitation of this review is the lack of well-designed, well-conducted studies addressing whether screening for speech and language delay or disorders improves outcomes.

CONCLUSIONS: Several screening tools can accurately identify children for diagnostic evaluations and interventions, but evidence is inadequate regarding applicability in primary care settings. Some treatments for young children identified with speech and language delays and disorders may be effective.

[a]*RTI International, Research Triangle Park, North Carolina; and* [b]*Division of Speech and Hearing Sciences, and Departments of* [c]*Pediatrics, and* [d]*Internal Medicine, School of Medicine, University of North Carolina, Chapel Hill, North Carolina*

Drs Wallace, Berkman, Watson, and Coyne-Beasley selected articles for inclusion; Drs Wallace, Berkman, Watson, Coyne-Beasley, and Wood extracted data; Drs Wallace, Berkman, Watson, and Wood and Ms Cullen drafted the manuscript; Drs Wallace, Berkman, Watson, Wood, and Lohr and Ms Cullen revised the manuscript; all authors reviewed the manuscript; and Dr Wallace gave final approval of the version to be published.

The staff at the Agency for Healthcare Research Quality and members of the US Preventive Services Task Force developed the scope of the work and reviewed draft manuscripts. Approval from the Agency for Healthcare Research Quality was required before the manuscript was submitted for publication, but the authors are solely responsible for the content and the decision to submit it for publication.

www.pediatrics.org/cgi/doi/10.1542/peds.2014-3889

DOI: 10.1542/peds.2014-3889

Accepted for publication Feb 11, 2015

Address Correspondence to Ina F. Wallace, Division for Health Services and Social Policy Research, RTI International, P.O. Box 12194, Research Triangle Park, NC 27709-2194. E-mail: wallace@rti.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Speech and language delays and disorders are common, with an estimated prevalence between 5% and 12% (median, 6%) in children 2 to 5 years of age.[1] A speech or language delay implies that the child is developing speech or language in the correct sequence but at a slower rate than expected, whereas a speech or language disorder suggests that the child's speech or language ability is qualitatively different from what is typical. In this review, we use speech and language "delay," "disorder," "impairment," and "disability" interchangeably.

The American Speech-Language-Hearing Association guidelines describe a speech disorder as an impairment of the articulation of speech sounds, fluency, or voice and a language disorder as impaired comprehension or use of spoken, written, or other symbol systems. A disorder may involve the form of language (phonology, morphology, syntax), the content of language (semantics), and the function of language in communication (pragmatics) in any combination.[2] Because prelinguistic communication behaviors (eg, gestures, babbling, joint attention) are associated with language delays,[3–5] this review considers screening of both verbal and preverbal communication skills.

Young children with speech and language delay in the preschool years may be at increased risk for learning disabilities once they reach school age.[6] Children with both speech sound disorders and language impairment are at greatest risk for language-based learning disabilities (eg, difficulties in reading and written language).[7,8] Estimates of the increased risk for poor reading outcomes in grade school are 4 to 5 times greater for children with speech and language impairment than for children with appropriate development[9–12]; risk persists into adulthood.[13] Adults who had speech and language disorders as children may hold lower-skilled jobs and are more likely to experience unemployment than other adults.[14] Behavior problems and impaired psychosocial adjustment associated with speech and language may also persist into adulthood.[15–17]

Identifying speech and language problems before children enter school can foster initiation of early interventions before these problems interfere with formal education and behavioral adjustment. AAP clinical guidelines recommend that pediatric health care providers perform surveillance at every well-child visit for children <36 months of age; should concerns arise, screening should be administered using standardized developmental tools.[18] Irrespective of concerns, the guidelines identify 9, 18, and 24 or 30 months as appropriate ages for developmental screening.

In 2006, the US Preventive Services Task Force (USPSTF) concluded that evidence was insufficient to recommend for or against ("I statement") routine use of brief, formal screening instruments in primary care to detect speech and language delay in children up to 5 years of age. In 2013, the USPSTF commissioned a new systematic review of the current evidence on brief, formal screening for speech and language delays and disorders in children 5 years old and younger.[19] The USPSTF used it to update its 2006 recommendations about screening in primary care settings.

METHODS

Following the USPSTF Procedure Manual,[20] we developed an analytic framework (Supplemental Fig 2), list of key questions (KQs), and supporting contextual questions. We searched Medline (via PubMed), the Cochrane Library, PsycInfo, and Cumulative Index to Nursing and Allied Health Literature for English-language articles published from January 1, 2004, through July 20, 2014. We conducted targeted searches for unpublished literature in ClinicalTrials.gov. Appendix A of the full report[19] documents the search strategy. To supplement electronic searches, we reviewed reference lists of pertinent review articles and included studies

We used a PICOTS (populations, interventions, comparators, outcomes, timing, settings, and study designs) approach to identify studies that met inclusion and exclusion criteria that we developed for each key question (see Appendices B and C of the full report).[19] Two reviewers independently applied inclusion and exclusion criteria to all studies in the 2006 review and to all new studies from our update searches.

An investigator abstracted evidence from included full-text articles for each key question; a second investigator checked and confirmed each abstraction. We also checked for errors in the abstractions of studies in the 2006 review. Two reviewers independently rated the quality of each study based on USPSTF guidelines as good, fair, or poor (see Appendix D of the full report);[19] they resolved discrepancies by discussion. We reassessed the quality rating of studies in the 2006 review to ensure that they met current criteria. If 1 reviewer disagreed with this earlier assessment, we rerated the quality of that study through dual review.

We abstracted accuracy statistics when available from screening studies. When investigators did not provide accuracy statistics, we calculated sensitivity, specificity, prevalence, positive and negative predictive values, positive and negative likelihood ratios (LRs), and 95% confidence intervals (CIs) for sensitivity and specificity (see Appendix E of the full report).[19,21]

We evaluated applicability to US primary care populations based on demographics, coexisting conditions,

representativeness of the population, study refusal rate, severity of the delay, and recruitment source and applicability of the intervention/screening (ie, how well the clinical experience is liable to be reproduced in other settings).

This review was funded by the Agency for Healthcare Research and Quality (AHRQ). The USPSTF members and AHRQ Medical Officers helped develop the scope, KQs, and analytic framework that guided our literature search and review.

RESULTS

We document the impact of screening using evidence derived from included studies identified through the 2006 report,[22,23] our database and manual searches,[19,24] and recommendations from peer reviewers. We had evidence for 5 of 7 KQs (Supplemental Fig 2); we had no evidence for KQ3 (adverse effects of screening) or KQ4 (surveillance by primary care clinicians). Figure 1 shows the flow of studies from initial identification of titles and abstracts to final inclusion or exclusion.

KQ1: Improvements in Outcomes

No study met the 2006 inclusion criteria to determine whether screening improved either speech and language or other outcomes. One randomized controlled trial (RCT) met our inclusion criteria by randomizing a large national sample of children who received regularly scheduled care at child health centers to early screening and measuring outcomes at 8 years of age.[25,26] We did not include evidence from this trial owing to a rating of poor quality caused by very high attrition.

KQ2: Accurate Identification of Children for Diagnostic Evaluations and Interventions: Screening Accuracy

We examined the accuracy of screening techniques and whether accuracy varies by demographic and screening source. We included 24 good and fair studies (26 articles):

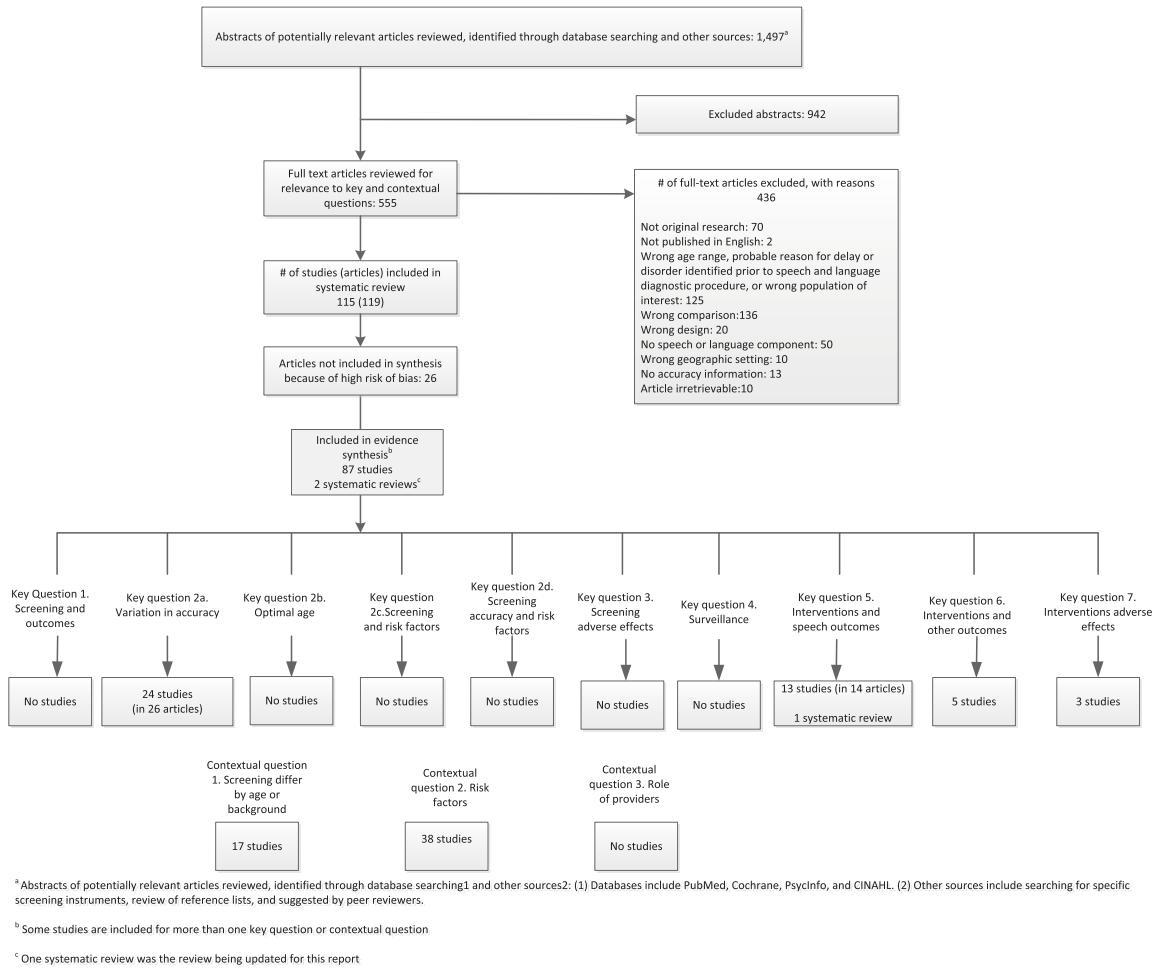


FIGURE 1

Flow diagram of study retrieval and selection. [a]Abstracts of potentially relevant articles reviewed, identified through database searching(1) and other sources(2): (1) Databases include PubMed, Cochrane, PsycInfo, and Cumulative Index to Nursing and Allied Health Literature. (2) Other sources include searching for specific screening instruments, review of reference lists, and suggested by peer reviewers. [b]Some studies are included for more than one key question or contextual question. [c]One systematic review was the review being updated for this report.

8 newly identified studies (9 articles[27–35]), and 16 studies (17 articles) from the 2006 review[36–52] (Supplemental Table 3). Supplemental Table 4 describes relevant screening instruments.

Detailed Synthesis of Evidence on Screening Accuracy

Tables 1 and 2 present accuracy statistics separately for parent and trained-examiner instruments, respectively. We report sensitivity and specificity (and 95% CIs), prevalence, positive and negative predictive values, and positive and negative LRs. We present median (not mean) values because accuracy statistics were skewed. We report the accuracy statistics by age group when possible.

Accuracy of Screening Instruments Used by Parents

Altogether, 14 studies (16 articles[27–30,32–35,40,42,43,46–49,52]) examined the accuracy of screeners in which parents rated the speech and language skills of their young children (mostly 2 or 3 years of age) (Table 1). Cutoff scores for positive screening (ie, a speech or language problem), when provided, varied by instrument.

Sensitivity for detecting a true speech and language delay or disorder using parent-report screeners ranged between 50% and 94% (median, 81%); specificity for detecting a child without speech and language delays ranged between 45% and 96% (median, 87%). Children with positive screening results (ie, those who failed the screening test) had a moderately[53] higher likelihood of language delay than children with negative screening results (ie, those who passed) in at least 1 study investigating the Ages and Stages Questionnaire (ASQ), the Communicative Development Inventory (CDI), the Language Development Survey (LDS), the Parent Questionnaire, and Ward's screening tool. With respect to negative LRs, results from ≥1 studies using the CDI, the Infant-Toddler Checklist (ITC), and LDS suggested a moderately lower likelihood of language delay for those children who passed the screening test relative to those who did not.

Accuracy by Age of Child

ASQ sensitivity was marginally higher for older children (4.5 years) in 1 study[27] than for younger children (2 to 3 years) in 2 other studies.[28,29] However, in the latter 2 studies, the positive LRs indicated at least a moderately higher likelihood of a language delay in children who screened positive relative to children who screened negative; we saw no such increase in the likelihood of delay in the study of older children. The negative LRs were small and equivalent for both younger and older samples.

Four of the 5 CDI studies examined the accuracy of the toddler version (18 to 36 months).[29,30,32–34] The fifth study used the preschool version with children 36 to 62 months of age.[28] Accuracy of the 2 versions was similar. The 1 ITC study separately considered 2 age groups of toddlers (12 to 17 months; 18 to 24 months); accuracy was similar for younger and older toddlers.[35]

Accuracy of Longer-Term Prediction

Two studies examined the accuracy of parent-reported screeners for predicting long-term language delay.[32,33,42,43] Both studies examined the accuracy of the screener at 2 years in relation to the reference standard (a diagnostic tool) at both 2 years and 3 years. In the LDS study,[43] sensitivity for detecting a language delay at 3 years was 67% (91% at 2 years). Specificity for detecting typical language development at 3 years was 93% (96% at 2 years). In the ELFRA-2 (ie, German CDI) study,[32,33] sensitivity and specificity at 3 years were 94% (93% at 2 years) and 61% (88% at 2 years), respectively.

Accuracy of Screening Instruments Used by Trained Examiners

Twelve studies examined the accuracy of instruments administered by trained examiners, including nurses, primary care providers, teachers, and paraprofessionals (Table 2).[27,31,36–39,41,44,45,48,50,51] These studies tended to focus on older preschool-age children: 3 studies included children 2 to 3 years of age[44,45,48]; 1 of children 3 to 4 years of age[37]; 5 of children 4 to 5 years of age[27,31,36,50,51]; and 3 of children across different ages (18 to 72 months).[38,39,41] Several studies included >1 screening instrument. All but 2 instruments require some direct testing of the child; the Developmental Nurse Screen[48] and the Davis Observation Checklist for Texas (DOCT)[36] involve ratings made after observing the child.

Sensitivity for detecting a true delay or disorder ranged between 17% and 100% (median, 74%); specificity for detecting typical speech and language ranged between 46% and 100% (median, 91%). In studies of the Battelle Developmental Inventory Screening Test,[27] DOCT,[36] Screening Kit of Language Development (SKOLD),[38] Sentence Repetition Screening Test,[51] Structured Screening Test,[44] and the Trial Speech Screening Test,[31] positive LRs indicated at least a moderately higher likelihood of language delay for those who screened positive; the studies of the Brigance Preschool Screening Test,[27] DOCT,[36] Early Screening Test,[27] Hackney Early Language Screening Test,[45] Northwestern Syntax Screening Test,[37] and SKOLD,[38] indicated at least a moderately lower likelihood of language delay for those who screened negative.

Accuracy by Age of Children and Language Dialect

One study used the SKOLD to screen children ages 30 to 48 months.[38] For versions appropriate for children 30

TABLE 1 Accuracy of screening instruments for speech and language delays and disorders: parent-rated screeners

Instrument and Version	Decision Cutoff Point	Reference	USPSTF Quality Rating	Child Age	*n*	Reference Instrument	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Prevalence, %[a]	PPV, %[a,b]	NPV, %[a,b]	PLR, %[a]	NLR, %[a]
Ages and Stages Questionnaire, 2nd ed.	"Recommended cutoff"	Frisk et al 2009[27]	Fair	4.5 y	110	PLS-4 Receptive	67 (45–88)	73 (64–82)	16	32	92	2.4	0.46
						PLS-4, Expressive	73 (54–91)	76 (67–85)	20	43	92	3.0	0.36
Ages and Stages Questionnaire, Spanish version	NR	Guiberson et al 2011[29]	Fair	24–35 mo	45	PLS-4, Spanish edition	56 (36–77)	95 (87–100)	51	92[c]	67[c]	12.4	0.46
Ages and Stages Questionnaire, Spanish version	NR	Guiberson and Rodríguez 2010[28]	Fair	32–36 mo	48	PLS-4, Spanish edition	59 (38–80)	92 (82–100)	46	87	73	7.7	0.44[c]
SCS18: Swedish CDI WS	<8 words	Westerlund et al 2006[34]	Fair	18 mo	891	Language Observation, 3 y	50 (34–66)	90 (88–92)	4	18[c]	89[c]	4.8[c]	0.56
CDI WS	<19th percentile	Heilmann et al 2005[30]	Fair	24 mo	100	PLS-3	81 (69–94)	79 (69–89)	38	70[c]	89[c]	3.9	0.23
ELFRA-2: German CDI Words and Sentences	<50 words or 50-80 words and grammatical scores below cutoff	Sachse and Von Suchodoletz 2008, 2009[32,33]	Good	24–26 mo	117	SETK-2	93 (87–99)	87 (78–97)	59	91[c]	89[c]	7.3	0.08
Short Form Inventarios del Desarrollo de Habilidades Comunicativas: Spanish CDI WS	NR	Guiberson et al 2011[29]	Fair	24–35 mo	45	PLS-4, Spanish edition	87 (73–100)	86 (72–100)	51	87[c]	86[c]	6.4	0.15
Pilot Inventario–III: Spanish CDI III	NR	Guiberson and Rodríguez 2010[28]	Fair	32–36 mo	48	PLS-4, Spanish edition	82 (66–98)	81 (66–96)	46	78	84	4.2[c]	0.22[c]
General Language Screen	≥2 of 11 items endorsed	Stott et al 2002[49]	Fair	36 mo	596	DPII (37 mo)	75 (67–83)	81 (77–84)	8[c]	47	94	3.9[c]	0.31[c]
						EAT, RDLS, BPVS (45 mo)	67 (—[d])	68 (—[d])	4[c]	31	91	—[d]	—[d]
Parent Language Checklist: previous version of the General Language Screen	1 failed item	Burden et al 1996[40]	Good	36 mo	425	Renfrew Action Picture Test, Bus Story, study-derived tests of phonology and comprehension	87 (82–93)	45 (39–51)	32	42	89	1.6	0.28
Infant-Toddler Checklist	NR	Wetherby et al 2003[35]	Fair	12–17 mo	151	CSBS Behavior Sample	89 (80–97)	74 (66–83)	35	65[c]	92[c]	3.5	0.15
				18–24 mo	81		86 (75–96)	77 (64–90)	52	80[c]	83[c]	3.7	0.19
Language Development Survey	<50 words or no word combinations	Klee et al 1998[42]	Fair	24–26 mo	64	Clinical judgment on infant MSEL language scales, MLU	91 (74–100)	87 (78–96)	17	59	98	6.9	0.10
	≥28 screening score	Klee et al 2000[e43]	Fair	24–26 mo	64		91 (74–100)	96 (91–100)	17	83	98	24.1	0.09
Language Development Survey Study 2	<50 words or no word combinations	Rescorla and Alley 2001[47]	Fair	25.4 mo	66	RDLS Expressive	94 (84–100)	67 (53–80)	27	52[c]	97[c]	2.8	0.08

TABLE 1 Continued

Instrument and Version	Decision Cutoff Point	Reference	USPSTF Quality Rating	Child Age	*n*	Reference Instrument	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Prevalence, %[a]	PPV, %[a,b]	NPV, %[a,b]	PLR, %[a]	NLR, %[a]
Language Development Survey Study 3	<50 words or no word combinations	Rescorla 1989[46]	Fair	24–34 mo	81	RDLS Expressive	89 (80–98)	86 (75–97)	56	89	86	6.4	0.13
Parent Questionnaire	NR	Stokes 1997[48]	Fair	34–40 mo	381	SLP rating using language sample, RDLS Comprehension	78 (66–89)	91 (88–94)	13	56[c]	96[c]	8.3	0.24
Ward's Created Screening Tool	≥1 item	Ward 1984[52]	Fair	7–23 mo	1070	REEL	80 (75–85)	92 (90–94)	24[c]	75	94	9.6	0.22

BPVS, British Picture Vocabulary Scales; CDI, Communicative Development Inventory; WS, Words and Sentences; CSBS, Communication and Symbolic Behavior Scales; DPII, Developmental Profile II; EAT, Edinburgh Articulation Test; ELFRA, Elternfragebogen fur die Fruberkennung von Riskokindern; MLU, mean length of utterances; MSEL, Mullen Scales of Early Learning; NR, not reported; NLR, negative likelihood ratio; NPV, negative predictive value; PLS, Preschool Language Scale; PLR, positive likelihood ratio; PPV, positive predictive value; REEL, Receptive Expressive Emergence of Language; RDLS, Reynell Developmental Language Scales; SCS18, Swedish Communication Screening at 18 mo of age; SETK-2, Sprachentwicklungstest fur sweijahrige slindes; SETK-3/5, Sprachentwicklungstest fur dreibis funfjahrige Kinder; SLP, speech language pathologist

[a] Calculated by EPC authors unless otherwise noted that study investigators provided data. Prevalence values were not estimated or weighted to reflect sampling

[b] Predictive values may be questionable for studies in which prevalence exceeded 10%; the problem arises when investigators choose a random sample of children with negative screens to complete the reference measures.

[c] Study investigators provided data.

[d] Could not calculate because of lack of data in article.

[e] Same data using a different decision rule for failing screener.

to 36 months, 37 to 42 months, and 43 to 48 months, median sensitivity rates were 94%, 94%, and 97%, respectively; median specificity rates were 92%, 88%, and 85%. Across the 3 age levels, median sensitivity and specificity were 88% and 86% for the African American dialect versions and 100% and 93% for the Standard English versions.

KQ5: Treatment: Speech and Language Outcomes

Thirteen RCTs (6 newly identified)[54–59] in 14 articles evaluating speech and language interventions and 1 systematic review met criteria for inclusion (Supplemental Table 5). Of these, 11 examined language outcomes and 8 measured speech outcomes. The systematic review of treatment of childhood apraxia of speech failed to find any studies that met our inclusion criteria, so we did not consider it further.[60]

Language

Of 11 studies measuring language outcomes (Supplemental Table 6), 4 used parents as the primary intervention agent.[57,61–63] In 2 trials testing variations of the Hanen Parent Program[57,62] for toddlers with language delays, 1 found significant effects on expressive language measures favoring the treatment group;[62] in contrast, another trial found no significant differences in receptive or expressive language.[57] Group training on language activities for parents of toddlers with limited expressive language found significant effects on expressive and receptive language.[61] Finally, 1 group of parents learned activities to target speech sounds and a second group of parent shared storybooks with their children[63]; neither treatment was associated with gains in child expressive syntax or semantic knowledge compared with the control group.

Two trials tested treatments primarily or exclusively delivered in

TABLE 2 Accuracy of Screening Instruments: Professional and Paraprofessional-Administered Screeners

Instrument and Component	Decision Cutoff Point	Reference	USPSTF Quality Rating	Child Age	*n*	Reference Instrument	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Prevalence[a]	PPV[a,b]	NPV[a,b]	PLR[a]	NLR[a]
Battelle Developmental Inventory Screening Test, Receptive	<1 SD	Frisk et al 2009[27]	Fair	4.5 y	110	PLS-4 Receptive	56 (33–78)	70 (60–79)	16	26	89	1.8	0.89
						PLS-4 Expressive	68 (49–88)	86 (79–94)	20	56	92	5.0	0.37
Brigance Preschool Screen	<1 SD	Frisk et al 2009[27]	Fair	4.5 y	110								
Receptive						PLS-4 Receptive	61 (39–84)	60 (50–70)	16	23	89	4.2	1.5
Expressive						PLS-4 Expressive	91 (79–100)	78 (70–87)	20	51	97	0.12	0.65
Davis Observation Checklist for Texas	NR	Alberts et al 1995[36]	Fair	52–67 mo	59	MSCA, GFTA, informal language sample	80 (55–100)	98 (94–100)	17	89	96	39.2	0.20
Denver Articulation Screening Test	<15th percentile	Drumwright et al 1973[41]	Fair	30–72 mo	150	Henja Articulation Test	92 (—[d])	97 (—[d])	—[d]	—[d]		—[d]	—[d]
Denver Developmental Screening Test Language Sector	NR	Borowitz and Glascoe 1986[39]	Fair	18–66 mo	71	PLS	46 (34–58)	100 (100–100)	92	100	15	—[e]	0.53
Developmental Nurse Screen	NR	Stokes 1997[48]	Fair	34–40 mo	378	SLP rating using language sample, RDLS Comprehension	76 (—[d])	96 (—[d])	—[d]	80	96	—[d]	—[d]
Early Screening Profile Verbal Concepts	<1 SD	Frisk et al 2009[27]	Fair	4.5 y	110	PLS-4 Auditory	94 (84–100)	68 (59–78)	16	40	98	3.0	0.08
						PLS-4 Expressive	86 (72–100)	81 (72–89)	20	53	96	4.5	0.17
Fluharty Preschool Screening Test	Failure ≥1 subtests	Allen and Bliss 1987[37]	Fair	36–47 mo	182	SICD	60 (41–79)	81 (75–87)	14	33	93	3.1	0.49
FPSLST Articulation	NR	Sturner et al 1993[50] study 1	Fair	53–68 mo	51	AAPS-R	74 (—[d])	96 (—[d])	4[c]	50	—[d]	—[d]	—[d]
FPSLST Language	NR	Sturner et al 1993[50] study 1	Fair	53–68 mo	51	TACL-R	38 (—[d])	85 (—[d])	17[c]	42	—[d]	—[d]	—[d]
FPSLST Articulation	NR	Sturner et al 1993[50] study 2	Fair	55–69 mo	147	TD	43 (—[d])	93 (—[d])	5[c]	26	—[c]	—[d]	—[d]
FPSLST Language	NR	Sturner et al 1993[50] study 2	Fair	55–69 mo	147	TOLD-P	17 (—[d])	97 (—[d])	22[c]	50	—[d]	—[d]	—[d]
Northwestern Syntax Screening Test	Failure ≥1 subtests	Allen and Bliss 1987[37]	Fair	36–47 mo	182	SICD	92 (81–100)	48 (41–56)	14	22	97	1.8	0.16
							81–100	41–56					
SKOLD		Bliss and Allen 1984[38]	Fair										
Standard English													
SKOLDS30	<11			30–36 mo	47	SICD	100 (100–100)	98 (93–100)	6	75	100	44.0	0
SKOLDS37	<10			37–42 mo	93	SICD	100 (100–100)	91 (85–97)	11	33	100	11.1	0
SKOLDS43	<19			43–48 mo	100	SICD	100 (100–100)	93 (88–98)	9	60	100	15.2	0
African American dialect													
SKOLDB30	<9			30–36 mo	75	SICD	89 (68–100)	86 (78–95)	12	47	98	6.5	0.13

TABLE 2 Continued

Instrument and Component	Decision Cutoff Point	Reference	USPSTF Quality Rating	Child Age	*n*	Reference Instrument	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Prevalence[a]	PPV[a,b]	NPV[a,b]	PLR[a]	NLR[a]
SKOLDB37	<14			37–42 mo	91	SICD	88 (65–100)	86 (78–92)	9	37	99	6.0	0.15
SKOLDB43	<19			43–48 mo	54	SICD	94 (84–100)	78 (64–91)	33	68	97	4.2	0.07
Sentence Repetition Screening Test	<20th percentile	Sturner et al 1996[51]	Fair	54–66 mo	323[i]	AAPS-R	57 (45–69)	95 (93–98)	19[c]	12.5	NR	0.45	NR
						ITPA, Bankson	62 (45–78)	91 (87–94)	11[c]	6.6	NR	0.42	NR
Structured Screening Test	<10	Laing et al 2002[44]		Good	282	RDLS	66 (54–77)	89 (85–94)	23	65	90	6.2[c]	0.38[c]
Hackney Early Language Screening Test, earlier version	≤10	Law 1994[45]	Good	30 mo	189	RDLS	98 (94–100)	69 (61–77)	26	53	99	3.17	0.03
Trial Speech Screening Test	<12 elements	Rigby and Chesham 1981[31]	Fair	54 mo	438	SLP evaluation of Renfrew, RDLS, Edinburgh Articulation	80 (68–92)	93 (91–96)	10	58	98	12.1	0.21

AAPS-R, Arizona Articulation Proficiency Scale: Revised; FPLST, Fluharty Preschool Speech and Language Screening Test; GFTA, Goldman-Fristoe Test of Articulation; ITPA, Illinois Test of Psycholinguistic Abilities; MSCA, McCarthy Scales of Children's Abilities; NR, not reported; PLS, Preschool Language Scale; RDLS, Reynell Developmental Language Scales; SICD, Sequenced Inventory of Communication Development; SKOLD, Screening Kit of Language Development; SLP, speech language pathologist; TACL-R, Test for Auditory Comprehension of Language – Revised; TD, Templin-Darley Tests of Articulation Consonant Singles Subtest; TOLD-P, Test of Language Development Primary.

[a] Calculated by EPC authors unless otherwise noted that study investigators provided data. Prevalence values were not estimated or weighted to reflect sampling

[b] Predictive values may be questionable for studies in which prevalence exceeded 10%; the problem arises when investigators choose a random sample of children with negative screens to complete the reference measures.

[c] Study investigators provided data.

[d] Could not calculate because of lack of data in article.

[e] Calculated as infinity.

a small group format by researcher-trained staff to toddlers[64] or preschoolers;[54] the latter also included individual treatment sessions after the first 10 weeks of the program.[54] Both trials reported significant improvement on measures of language skills.

Four trials tested individual treatment to children by research staff or speech-language pathologists.[58,59,65,66] One examined the effects of providing young children (18 to 42 months) with language or phonological delays with access to usual speech-language therapy services in the community.[65] With an average of only 6.2 hours of therapy over 12 months, children showed small but significant gains in receptive, but not expressive, language relative to controls. Another trial involving 4-year-olds with specific language impairments tested a manualized intervention that addressed individualized language goals, phonological and print awareness, and letter knowledge.[59] The intervention had no significant effect on expressive, receptive, or pragmatic language. A third trial tested the effects of a strategy called recasting (repeating what is said by a child, with correct articulation or with a grammatical expansion of the child's utterance).[58] The intervention had no overall effect on children's mean length of utterances but did produce improvements among children with the lowest baseline articulation skills. The fourth trial tested whether an individualized treatment of children with speech sound disorders affected mean length of utterance but found no significant language effect.[66]

Finally, preschoolers with language impairments who played with peers with age-appropriate language skills in the house play area of their classroom over a 3-week period improved significantly on activity-specific expressive language.[64]

Speech Sounds

Eight trials reported on various speech sounds[54,58,59,63,65–68] (Supplemental Table 6). Of 2 trials of parent-mediated interventions, 1 found that a modified Hanen Parent Program had significant effects on consonant inventory and syllable structure.[67] In the other trial, parents engaged the child in activities directed at discrimination of sounds.[63] Children in the control condition improved more in auditory discrimination in the presence of background noise than experimental subjects.

A small group intervention for toddlers significantly improved the percentage of intelligible utterances for treated children.[68]

Two studies examined individual treatment by speech-language pathologists. One examined the effects of the "cycles" approach to phonological therapy (wherein rule-based errors in the child's speech sound production are treated through recursive cycles of therapy) for preschoolers with severe phonological disorders; the intervention produced significant effects on standardized tests and percentage of correct consonants from a speech sample.[66] The other study found no improvement in phonology error rate for children randomized to usual community speech-language pathology services for a year; however, treated children were 2.7 times less likely to exhibit the severity of speech sound problems used as a criterion for initial study eligibility.[65]

The recasting trial found no main effects on children's intelligibility, but did find improvements among children with the lowest baseline articulation skills.[58]

Two studies reported that their interventions significantly improved phonological awareness skills in preschoolers. In one, teaching assistants delivered small group and individual lessons[54]; in the other, language assistants provided individual home-based interventions.[59]

Fluency

Two trials examined the Lidcombe Program of Early Stuttering intervention.[69] Both significantly reduced stuttering in preschoolers, when delivered in a clinic setting[55] and when using a telephone-based health delivery model.[56]

KQ6: Treatment: Outcomes Other Than Speech and Language

Two trials examined effects on socialization. One, among children receiving community-based speech-language services, produced no significant effect.[65] The other, among language-delayed toddlers receiving small-group therapy, produced large and significant differences favoring the treated children.[68]

For reducing behavior problems, one trial tested the effectiveness of a low-intensity parent group program[57] and another an in-home individualized program provided by a language assistant[59]; neither found significant effects. Similarly, measures of well-being of toddlers[65] and health-related quality of life of preschoolers[59] yielded nonsignificant effects of treatment.

In 2 trials, toddlers randomized to speech-language services were no different from controls on attention level or play.[65,68] Parents of language-delayed toddlers participating in small-group language therapy reported significantly greater improvements in parental stress.[68]

Two trials measuring emergent literacy skills among preschoolers[54,59] found that letter knowledge improved significantly, but one failed to find a significant effect for a broader construct of literacy.[54] However, treatment did significantly improve a measure of reading comprehension administered at 6-months of follow-up.

KQ7. Adverse Effects of Treatments

Three studies examined potential adverse effects of interventions but reported no negative impacts on children or parents.[59,65,68]

DISCUSSION

Screening Accuracy

Some screening instruments accurately identify children for language delays or disorders. As in the 2006 review, however, we observed wide ranges of reported sensitivity and specificity; no one instrument clearly demonstrated the best characteristics or 1 age as optimal for screening. We compared findings from the same instrument in different populations; specifically, accuracy of 3 parent-rated screeners (ASQ, CDI, and ITC) and 2 trained-examiner screeners (Fluharty Preschool Speech and Language Screening Test and SKOLD) across ages. CDI, ITC, and SKOLD displayed consistency and acceptable levels of sensitivity and specificity ($\geq$70%)[70] at each age level; this suggests that they are more robust across different ages than ASQ and Fluharty Preschool Speech and Language Screening Test, which had generally low sensitivity across age levels.

Accuracy apparently drops over time. In the 2 studies[32,33,42,43] that examined whether a parent-report screener administered at 2 years would be as accurate at 3 years, sensitivity was lower in 1 study and specificity was lower in the other. Decreasing specificity with time may mean some that some children with language delays will "catch up" and display more typical language skills as they age.[71]

The comparison between parent-rated and trained-examiner screeners indicated many similarities in performance characteristics. Aside from the Denver Developmental Screening Test (now known as the Denver II), most trained-examiner tools are not used in primary care

offices and would require a dedicated, trained examiner to test the child directly. Three parent-rated screeners (CDI, ITC, and LDS) display acceptable sensitivity and specificity. Moreover, because parents complete these screeners, adopting them in a screening program would not burden a primary care practice with training someone in test administration. The more extensive information that parents provide related specifically to their children's language skills may help explain their greater accuracy in identifying children with speech and language delays than broad-based screeners that include other domains but fewer speech and language items. Moreover, staff in primary care settings could likely interpret results from parent screeners with little difficulty.

Treatment Outcomes

The majority of the 13 trials support the effectiveness of treating young children with language delays and disorders (6 of 11 trials reporting significant positive results) and those with problems with speech sounds (6 of 8 trials reporting significant positive results), and toddlers and preschoolers for fluency problems (2 of 2 trials reporting significant positive results). Individual and small-group service delivery models and various intervention agents, including parents supported or trained by professionals, speech-language pathologists, and trained teaching or therapy assistants, generally favored intervention groups.

Multiple factors limit the confident interpretation of this body of evidence on speech and language treatment. These factors involve (1) the small size of many trials, which constrains investigating moderators and mediators of treatment effectiveness; (2) the lack of replicated positive findings for any treatment approach except the Lidcombe program for stuttering; (3) the wide variability across trials in the age of children treated, intervention agents (eg, speech-language pathologists, teaching assistants, parents, research staff), intensity, content, and strategies; (4) the relatively small number of trials using manualized treatments or providing enough details of the treatment to permit replication; (5) a corresponding lack of data detailing treatment fidelity in many trials; (6) a lack of common outcome measures; and (7) inconsistency in how results are reported. Because of this degree of heterogeneity, we could not do any meta-analysis. Overall, the evidence offers little guidance about specific factors associated with effective treatments for young children with speech and language delays.

Contextual Issues: Risk Factors

One contextual issue involved whether consistent, reliable, and valid risk factors exist that clinicians could use to identify children at highest risk for speech and language delay and disorders.[19] We examined 31 cohort studies, 24 with multivariate analysis to control for other factors, and 1 review of studies on characteristics of late-talking toddlers; 20 cohort studies involved English-speaking children (Supplemental Tables 7 and 8). Potential risk factors for speech and language problems include male gender, family history of speech or language impairment, lower levels of parental education, and various perinatal risk factors (eg, prematurity, birth difficulties, and low birth weight).

Studies about risk factors varied in the type of delay or disorder being considered, used inconsistent measurement of risk factors, included heterogeneous patient populations, and inconsistently adjusted for confounders in multivariate models. Future research should account for the heterogeneity across populations of children, consider a multifactorial perspective of child development, examine social determinants of health as possible risk factors, and adopt more standardized outcome measures over a longer-term period of follow-up than has been customary to date.

Limitations of the Review

Numerous limitations of the literature base continue to plague the field. Some date to the 2006 review, but additional limitations we encountered further reduce the applicability of the findings.

Most serious is the lack of well-designed, well-conducted studies addressing the overarching question: does screening for speech and language delay or disorders improve outcomes? Moreover, neither the 2006 review nor our update found any studies that addressed the questions of adverse effects of screening or the role of enhanced surveillance by primary care clinicians in accurately identifying children for diagnostic evaluations and interventions, 2 important issues in screening.

We identified some instruments that can accurately screen children with speech and language delays. However, many studies included potentially inappropriate populations, such as "samples" of children identified (randomly or otherwise) on the basis of their language status. Using such "predetermined" samples hampers investigators from determining certain accuracy statistics (other than sensitivity and specificity) and may bias conclusions about screening accuracy and, thus, can limit applicability to pediatric populations in general. Moreover, few studies examined how well screeners detected speech and language disorders over the long term. Such studies are critical in calculating the real benefit of early detection. In addition, few of the screening accuracy studies occurred in primary care settings, and none in the United States. The extent to which conclusions reached from screening in primary care settings in Sweden, Australia, and the United Kingdom are generalizable to the United States is not known.

Most treatment studies were also conducted outside the United States. Whether conclusions reached from trials in countries with different medical, health insurance, and educational systems apply in this country remains an open question. Additional limitations relate to interpreting treatment outcomes and replicating interventions. Much of the literature lacks information about important features of the intervention, such as whether children received community services for speech and language outside the study, and does not adequately document intervention models. Finally, control groups in numerous trials were children offered intervention on a delayed schedule. This condition likely would make parents more willing to consent to enrolling their children in a RCT, but it constrains our ability to look at long-range outcomes for treated versus untreated children.

Future Research Needs

To determine whether screening for speech and language delay or disorders improves speech, language, or other outcomes, studies need to be specifically designed and executed to examine these issues. Furthermore, they need to be implemented with little risk of bias. This research gap presents an opportunity for a large study in primary care settings to test the efficacy of systematic routine screening for speech and language delays and disorders in comparison with not implementing routine screening. In tandem with this, the field would benefit from a study to examine the feasibility of speech and language-specific screening as part of the more general developmental screening that is already recommended.[18]

Given federal mandates under the Individuals with Disabilities Education Act that all children with a documented speech or language delay receive early intervention, conducting RCTs to examine the efficacy of interventions may be difficult in future. Protocols may adopt rigorous quasi-experimental designs, such as regression discontinuity designs, to answer intervention questions. Well-designed and implemented regression discontinuity designs meet standards for rigor for evaluations of evidence sponsored by the Institute of Education Sciences.

We recommend that stakeholders with an interest in screening develop research agendas and funding targeted to answer the important questions that we could not address. Future systematic reviews will benefit from an enhanced literature base.

CONCLUSIONS

We found no evidence to answer the overarching question of whether screening for speech and language delay or disorders improves speech and language outcomes. Studies from the 2006 review and our newly identified studies suggest that some screening instruments can accurately pinpoint these disorders. Although the parent-rated instruments require only that the primary care provider interpret the findings, studies have not examined this in practice. As in the 2006 review, we found no studies that addressed the harms of screening for speech and language delays. Neither did we find any evidence about the role of enhanced surveillance by a primary care clinician once a child elicits clinical concern for speech and language delay. Building on the 2006 review, we found evidence supporting the effectiveness of treating speech and language delays and disorders in children. Nevertheless, the whole body of evidence does not provide guidance regarding specific factors associated with effective treatments for young children with speech and language delays or disorders. Finally, we found no evidence relating to the harms of treating speech and language delays or disorders.

ACKNOWLEDGMENTS

This work was supported by the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services contract no. HHSA290201200015iTO2. The views expressed in this manuscript do not represent and should not be construed to represent a determination or policy of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

The authors gratefully acknowledge the following individuals for their contributions to this project and deeply appreciate their considerable support, commitment, and contributions: our AHRQ Medical Officer, Karen C. Lee, MD, MPH; the USPSTF leads; Evidence-based Practice Center (EPC) Project Manager, Carol Woodell, BSPH; RTI-UNC EPC Director, Meera Viswanathan, PhD; our EPC Librarian, B Lynn Whitener, MSLS; and our EPC publications specialist, Loraine Monroe.

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services Contract No. HHSA290201200015iTO2. The views expressed in this manuscript do not represent and should not be construed to represent a determination or policy of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

COMPANION PAPER: A companion to this article can be found on page XXX, and online at www.pediatrics.org/cgi/doi/10.1542/peds.2015-0211.

REFERENCES

1. Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord*. 2000;35(2):165–188
2. American Speech-Language-Hearing Association. Definitions of Communication Disorders and Variations. Available at: http://www.asha.org/policy/RP1993-00208/. Accessed February 23, 2015
3. Whitehurst GJ, Smith M, Fischel JE, Arnold DS, Lonigan CJ. The continuity of babble and speech in children with specific expressive language delay. *J Speech Hear Res*. 1991;34(5):1121–1129
4. Thal DJ, Tobias S. Communicative gestures in children with delayed onset of oral expressive vocabulary. *J Speech Hear Res*. 1992;35(6):1281–1289
5. Crais ER, Watson LR, Baranek GT. Use of gesture development in profiling children's prelinguistic communication skills. *Am J Speech Lang Pathol*. 2009; 18(1):95–108
6. Bashir AS, Scavuzzo A. Children with language disorders: natural history and academic success. *J Learn Disabil*. 1992;25(1):53–65, discussion 66–70
7. Raitano NA, Pennington BF, Tunick RA, Boada R, Shriberg LD. Pre-literacy skills of subgroups of children with speech sound disorders. *J Child Psychol Psychiatry*. 2004;45(4):821–835
8. Peterson RL, Pennington BF, Shriberg LD, Boada R. What influences literacy outcome in children with speech sound disorder? *J Speech Lang Hear Res*. 2009;52(5):1175–1188
9. Catts HW, Fey ME, Tomblin JB, Zhang X. A longitudinal investigation of reading outcomes in children with language impairments. *J Speech Lang Hear Res*. 2002;45(6):1142–1157
10. Catts H, Fey M, Zhang X, et al. Estimating the risk of future reading difficulties in kindergarten children: a research-based model and its clinical implementation. *Lang Speech Hear Serv Sch*. 2001;32(1):38–50
11. Tomblin JB, Zhang X, Buckwalter P, Catts H. The association of reading disability, behavioral disorders, and language impairment among second-grade children. *J Child Psychol Psychiatry*. 2000;41(4):473–482
12. Glogowska M, Roulstone S, Peters TJ, Enderby P. Early speech- and language-impaired children: linguistic, literacy, and social outcomes. *Dev Med Child Neurol*. 2006;48(6):489–494
13. Young AR, Beitchman JH, Johnson C, et al. Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *J Child Psychol Psychiatry*. 2002;43(5):635–645
14. Felsenfeld S, Broen PA, McGue M. A 28-year follow-up of adults with a history of moderate phonological disorder: educational and occupational results. *J Speech Hear Res*. 1994;37(6): 1341–1353
15. Law J, Rush R, Schoon I, Parsons S. Modeling developmental language difficulties from school entry into adulthood: literacy, mental health, and employment outcomes. *J Speech Lang Hear Res*. 2009;52(6):1401–1416
16. Cohen NJ, Barwick MA, Horodezky NB, Vallance DD, Im N. Language, achievement, and cognitive processing in psychiatrically disturbed children with previously identified and unsuspected language impairments. *J Child Psychol Psychiatry*. 1998;39(6): 865–877
17. Cohen NJ, Menna R, Vallance DD, Barwick MA, Im N, Horodezky NB. Language, social cognitive processing, and behavioral characteristics of psychiatrically disturbed children with previously identified and unsuspected language impairments. *J Child Psychol Psychiatry*. 1998;39(6):853–864
18. Council on Children With Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. *Pediatrics*. 2006;118(1):405–420. Available at: www.pediatrics.org/cgi/content/full/118/1/e405
19. Berkman ND, Wallace IF, Watson L, et al. Screening for Speech and Language Delay and Disorders in Children Age 5 Years or Younger: A Systematic Evidence Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 120. AHRQ Publication No. 13-05197-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
20. U.S. Preventive Services Task Force. Procedure Manual. Rockville, MD: USPSTF Program Office; 2011 August. www.uspreventiveservicestaskforce.org/Page/Name/procedure-manual. Accessed October 2014
21. Hamm RM. *Clinical Decision Making Calculators*. Oklahoma City: Department of Family and Preventive Medicine, The University of Oklahoma Health Sciences Center; 2004. Available at http://www.fammed.ouhsc.edu/robhamm/cdmcalc.htm2014. Accessed at http://omerad.msu.edu/ebm/index.html2014. Accessed September 14, 2014
22. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics*. 2006;117(2):e298–e319. Available at: www.pediatrics.org/cgi/content/full/117/2/e298
23. Nelson HD, Nygren P, Walker M, Penoscha R. Screening for speech and language delay in preschool children. Rockville, MD: Agency for Healthcare Research and Quality; February 2006
24. Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for speech and language delay: a systematic review of the literature. *Health Technol Assess*. 1998; 2(9):1–184
25. de Koning HJ, de Ridder-Sluiter JG, van Agt HME, et al. A cluster-randomised trial of screening for language disorders in toddlers. *J Med Screen*. 2004;11(3):109–116
26. van Agt HM, van der Stege HA, de Ridder-Sluiter H, Verhoeven LT, de Koning HJ. A cluster-randomized trial of screening for language delay in toddlers: effects on school performance and language development at age 8. *Pediatrics*. 2007;120(6):1317–1325. Available at: www.pediatrics.org/cgi/content/full/120/6/e1317
27. Frisk V, Montgomery L, Boychyn E, et al. Why screening Canadian preschoolers

for language delays is more difficult than it should be. *Infants Young Child*. 2009;22(4):290–308

28. Guiberson M, Rodríguez BL. Measurement properties and classification accuracy of two spanish parent surveys of language development for preschool-age children. *Am J Speech Lang Pathol*. 2010;19(3):225–237
29. Guiberson M, Rodríguez BL, Dale PS. Classification accuracy of brief parent report measures of language development in Spanish-speaking toddlers. *Lang Speech Hear Serv Sch*. 2011;42(4):536–549
30. Heilmann J, Ellis Weismer S, Evans J, Hollar C. Utility of the MacArthur-Bates communicative development inventory in identifying language abilities of late-talking and typically developing toddlers. *Am J Speech Lang Pathol*. 2005;14(1):40–51
31. Rigby MJ, Chesham I. A trial speech screening test for school entrants. *Br Med J (Clin Res Ed)*. 1981;282(6262):449–451
32. Sachse S, Von Suchodoletz W. Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr*. 2008;29(1):34–41
33. Sachse S, Von Suchodoletz W. [Response]. *J Dev Behav Pediatr*. 2009;30(2):176
34. Westerlund M, Berglund E, Eriksson M. Can severely language delayed 3-year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur-Bates Communicative Development Inventories. *J Speech Lang Hear Res*. 2006;49(2):237–247
35. Wetherby AM, Goldstein H, Clearly J, et al. Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants Young Child*. 2003;16(2):161–174
36. Alberts FM, Davis BL, Prentice L. Validity of an observation screening instrument in a multicultural population. *J Early Interv*. 1995;19(2):168–177
37. Allen DV, Bliss LS. Concurrent validity of two language screening tests. *J Commun Disord*. 1987;20(4):305–317
38. Bliss LS, Allen DV. Screening Kit of Language Development: a preschool language screening instrument. *J Commun Disord*. 1984;17(2):133–141
39. Borowitz KC, Glascoe FP. Sensitivity of the Denver Developmental Screening Test in speech and language screening. *Pediatrics*. 1986;78(6):1075–1078. Available at: www.pediatrics.org/cgi/content/full/78/6/e1075
40. Burden V, Stott CM, Forge J, Goodyer I. The Cambridge Language and Speech Project (CLASP). I. Detection of language difficulties at 36 to 39 months. *Dev Med Child Neurol*. 1996;38(7):613–631
41. Drumwright A, Van Natta P, Camp B, Frankenburg W, Drexler H. The Denver articulation screening exam. *J Speech Hear Disord*. 1973;38(1):3–14
42. Klee T, Carson DK, Gavin WJ, Hall L, Kent A, Reece S. Concurrent and predictive validity of an early language screening program. *J Speech Lang Hear Res*. 1998;41(3):627–641
43. Klee T, Pearce K, Carson DK. Improving the positive predictive value of screening for developmental language disorder. *J Speech Lang Hear Res*. 2000;43(4):821–833
44. Laing GJ, Law J, Levin A, Logan S. Evaluation of a structured test and a parent led method for screening for speech and language problems: prospective population based study. *BMJ*. 2002;325(7373):1152
45. Law J. Early language screening in city and Hackney: the concurrent validity of a measure designed for use with 2 1/2-year-olds. *Child Care Health Dev*. 1994;20(5):295–308
46. Rescorla L. The Language Development Survey: a screening tool for delayed language in toddlers. *J Speech Hear Disord*. 1989;54(4):587–599
47. Rescorla L, Alley A. Validation of the language development survey (LDS): a parent report tool for identifying language delay in toddlers. *J Speech Lang Hear Res*. 2001;44(2):434–445
48. Stokes SF. Secondary prevention of paediatric language disability: a comparison of parents and nurses as screening agents. *Eur J Disord Commun*. 1997;32(2 Spec No):139–158.
49. Stott CM, Merricks MJ, Bolton PF, Goodyer IM. Screening for speech and language disorders: the reliability, validity and accuracy of the General Language Screen. *Int J Lang Commun Disord*. 2002;37(2):133–151
50. Sturner RA, Heller JH, Funk SG, Layton TL. The Fluharty Preschool Speech and Language Screening Test: a population-based validation study using sample-independent decision rules. *J Speech Hear Res*. 1993;36(4):738–745
51. Sturner RA, Funk SG, Green JA. Preschool speech and language screening: further validation of the sentence repetition screening test. *J Dev Behav Pediatr*. 1996;17(6):405–413
52. Ward S. Detecting abnormal auditory behaviours in infancy: the relationship between such behaviours and linguistic development. *Br J Disord Commun*. 1984;19(3):237–251
53. Ebell M, Barry H. *Likelihood Ratios Part 1: Introduction*. Lansing, MI: Office of Medial Education Research and Development, College of Human Medicine, Michigan State University; 2008. Available at: http://omerad.msu.edu/ebm/index.html. Accessed September 18, 2014.
54. Fricke S, Bowyer-Crane C, Haley AJ, Hulme C, Snowling MJ. Efficacy of language intervention in the early years. *J Child Psychol Psychiatry*. 2013;54(3):280–290
55. Jones M, Onslow M, Packman A, et al. Randomised controlled trial of the Lidcombe programme of early stuttering intervention. *BMJ*. 2005;331(7518):659
56. Lewis C, Packman A, Onslow M, Simpson JM, Jones M. A phase II trial of telehealth delivery of the Lidcombe Program of Early Stuttering Intervention. *Am J Speech Lang Pathol*. 2008;17(2):139–149
57. Wake M, Tobin S, Girolametto L, et al. Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's Learn Language cluster randomised controlled trial. *BMJ*. 2011;343:d4741
58. Yoder P, Camarata M, Gardner E. Treatment effects on speech intelligibility and length of utterance in children with specific language and intelligibility impairments. *J Early Interv*. 2005;28(1):34–49

59. Wake M, Tobin S, Levickis P, et al. Randomized trial of a population-based, home-delivered intervention for preschool language delay. *Pediatrics*. 2013;132(4):e895–e904. Available at: www.pediatrics.org/cgi/content/full/132/4/e895

60. Morgan AT, Vogel AP. A Cochrane review of treatment for childhood apraxia of speech. *Eur J Phys Rehabil Med*. 2009; 45(1):103–110

61. Gibbard D. Parental-based intervention with pre-school language-delayed children. *Eur J Disord Commun*. 1994; 29(2):131–150

62. Girolametto L, Pearce PS, Weitzman E. Interactive focused stimulation for toddlers with expressive vocabulary delays. *J Speech Hear Res*. 1996;39(6): 1274–1283

63. Shelton RL, Johnson AF, Ruscello DM, Arndt WB. Assessment of parent-administered listening training for preschool children with articulation deficits. *J Speech Hear Disord*. 1978; 43(2):242–254

64. Robertson SB, Ellis Weismer S. The influence of peer models on the play scripts of children with specific language impairment. *J Speech Lang Hear Res*. 1997;40(1):49–61

65. Glogowska M, Roulstone S, Enderby P, Peters TJ. Randomised controlled trial of community based speech and language therapy in preschool children. *BMJ*. 2000;321(7266):923–926

66. Almost D, Rosenbaum P. Effectiveness of speech intervention for phonological disorders: a randomized controlled trial. *Dev Med Child Neurol*. 1998;40(5): 319–325

67. Girolametto L, Pearce PS, Weitzman E. Effects of lexical intervention on the phonology of late talkers. *J Speech Lang Hear Res*. 1997;40(2):338–348

68. Robertson SB, Ellis Weismer S. Effects of treatment on linguistic and social skills in toddlers with delayed language development. *J Speech Lang Hear Res*. 1999;42(5):1234–1248

69. Onslow M, Packman A, Harrison E, eds. *The Lidcombe Program of Early Stuttering Intervention: A Clinician's Guide*. Austin, TX: Pro-Ed; 2003

70. Macias MM. Developmental Screening Tools, D-PIP Training Workshop. Elk Grove Village, IL: American Academy of Pediatrics; 2006

71. Rescorla L. Late talkers: do good predictors of outcome exist? *Dev Disabil Res Rev*. 2011;17(2):141–150

72. Squires J, Potter L, Bricker D. *Ages and Stages Questionnaire user's guide*. Baltimore, MD: Brookes; 1999

73. Sturner RS, Layton TL, Evans AW, et al. Preschool speech and language screening: A review of currently available tests. *Am J Speech Lang Pathol*. 1994;3:25–36

74. Fenson L, Marchman VA, Thal DJ, et al. *MacArthur-Bates Communicative Development Inventories: Users guide and technical manual*. Baltimore, MD: Brookes; 2007

75. Jackson-Maldonado D, Marchman VA, Fernald LCH. Short-form versions of the Spanish MacArthur–Bates Communicative Development Inventories. *Appl Psycholinguist*. 2013; 34(4):837–868.

76. Sturner RA, Kunze L, Funk SG, Green JA. Elicited imitation: its effectiveness for speech and language screening. *Dev Med Child Neurol*. 1993;35(8):715–726

77. Adams-Chapman I, Bann CM, Vaucher YE, et al. Association between feeding difficulties and language delay in preterm infants using Bayley Scales of Infant Development-Third Edition. *J Pediatr*. 2013;163(3):680–685

78. Alston E, James-Roberts IS. Home environments of 10-month-old infants selected by the WILSTAAR screen for pre-language difficulties. *Int J Lang Commun Disord*. 2005;40(2):123–136

79. Campbell TF, Dollaghan CA, Rockette HE, et al. Risk factors for speech delay of unknown origin in 3-year-old children. *Child Dev*. 2003;74(2):346–357

80. Choudhury N, Benasich AA. A family aggregation study: the influence of family history and other risk factors on language development. *J Speech Lang Hear Res*. 2003;46(2):261–272

81. Desmarais C, Sylvestre A, Meyer F, Bairati I, Rouleau N. Systematic review of the literature on characteristics of late-talking toddlers. *Int J Lang Commun Disord*. 2008;43(4):361–389

82. Everitt A, Hannaford P, Conti-Ramsden G. Markers for persistent specific expressive language delay in 3-4-year-olds. *Int J Lang Commun Disord*. 2013; 48(5):534–553

83. Foster-Cohen SH, Friesen MD, Champion PR, Woodward LJ. High prevalence/low severity language delay in preschool children born very preterm. *J Dev Behav Pediatr*. 2010; 31(8):658–667

84. Fox AV, Dodd B, Howard D. Risk factors for speech disorders in children. *Int J Lang Commun Disord*. 2002;37(2): 117–131

85. Glascoe FP, Leew S. Parenting behaviors, perceptions, and psychosocial risk: impacts on young children's development. *Pediatrics*. 2010;125(2):313–319. Available at: www.pediatrics.org/cgi/content/full/125/2/e313

86. Hammer CS, Farkas G, Maczuga S. The language and literacy development of Head Start children: a study using the Family and Child Experiences Survey database. *Lang Speech Hear Serv Sch*. 2010;41(1):70–83

87. Harrison LJ, McLeod S. Risk and protective factors associated with speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *J Speech Lang Hear Res*. 2010;53(2):508–529

88. Henrichs J, Rescorla L, Schenk JJ, et al. Examining continuity of early expressive vocabulary development: the generation R study. *J Speech Lang Hear Res*. 2011;54(3):854–869

89. Kerstjens JM, de Winter AF, Bocca-Tjeertes IF, ten Vergert EM, Reijneveld SA, Bos AF. Developmental delay in moderately preterm-born children at school entry. *J Pediatr*. 2011;159(1): 92–98

90. Kerstjens JM, Bos AF, ten Vergert EMJ, de Meer G, Butcher PR, Reijneveld SA. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Hum Dev*. 2009;85(7):443–447

91. Kerstjens JM, de Winter AF, Bocca-Tjeertes IF, Bos AF, Reijneveld SA. Risk of developmental delay increases exponentially as gestational age of preterm infants decreases: a cohort study at age 4 years. *Dev Med Child Neurol*. 2012;54(12):1096–1101

92. Law J, Rush R, Anandan C, Cox M, Wood R. Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*. 2012; 130(1):e132–e137. Available at: www.pediatrics.org/cgi/content/full/130/1/e132

93. Mossabeb R, Wade KC, Finnegan K, Sivieri E, Abbasi S. Language development survey provides a useful screening tool for language delay in preterm infants. *Clin Pediatr (Phila)*. 2012;51(7):638–644

94. O'Leary C, Zubrick SR, Taylor CL, Dixon G, Bower C. Prenatal alcohol exposure and language delay in 2-year-old children: the importance of dose and timing on risk. *Pediatrics*. 2009;123(2):547–554. Available at: www.pediatrics.org/cgi/content/full/123/2/e547

95. Peña ED, Gillam RB, Bedore LM, Bohman TM. Risk for poor performance on a language screening measure for bilingual preschoolers and kindergarteners. *Am J Speech Lang Pathol*. 2011;20(4):302–314

96. Potijk MR, Kerstjens JM, Bos AF, Reijneveld SA, de Winter AF. Developmental delay in moderately preterm-born children with low socioeconomic status: risks multiply. *J Pediatr*. 2013;163(5):1289–1295

97. Pruitt SL, Garrity AW, Oetting JB. Family history of speech and language impairment in African American children: implications for assessment. *Top Lang Disord*. 2010;30(2):154–164

98. Reilly S, Wake M, Bavin EL, et al. Predicting language at 2 years of age: a prospective community study. *Pediatrics*. 2007;120(6):e1441–e1449. Available at: www.pediatrics.org/cgi/content/full/120/6/e1441

99. Reilly S, Onslow M, Packman A, et al. Predicting stuttering onset by the age of 3 years: a prospective, community cohort study. *Pediatrics*. 2009;123(1): 270–277. Available at: www.pediatrics.org/cgi/content/full/123/1/e270

100. Reilly S, Onslow M, Packman A, et al. Natural history of stuttering to 4 years of age: a prospective community-based study. *Pediatrics*. 2013;132(3):460–467. Available at: www.pediatrics.org/cgi/content/full/132/3/e460

101. Roth C, Magnus P, Schjølberg S, et al. Folic acid supplements in pregnancy and severe language delay in children. *JAMA*. 2011;306(14):1566–1573

102. Schjølberg S, Eadie P, Zachrisson HD, Oyen AS, Prior M. Predicting language development at age 18 months: data from the Norwegian Mother and Child Cohort Study. *J Dev Behav Pediatr*. 2011; 32(5):375–383

103. Singer LT, Siegel AC, Lewis B, Hawkins S, Yamashita T, Baley J. Preschool language outcomes of children with history of bronchopulmonary dysplasia and very low birth weight. *J Dev Behav Pediatr*. 2001;22(1):19–26

104. Tallal P, Ross R, Curtiss S. Familial aggregation in specific language impairment. *J Speech Hear Disord*. 1989;54(2):167–173

105. Tomblin JB, Hardy JC, Hein HA. Predicting poor-communication status in preschool children using risk factors present at birth. *J Speech Hear Res*. 1991;34(5):1096–1105

106. Tomblin JB, Smith E, Zhang X. Epidemiology of specific language impairment: prenatal and perinatal risk factors. *J Commun Disord*. 1997;30(4): 325–343, quiz 343–344

107. van Batenburg-Eddes T, Henrichs J, Schenk JJ, et al. Early infant neuromotor assessment is associated with language and nonverbal cognitive function in toddlers: the Generation R Study. *J Dev Behav Pediatr*. 2013;34(5): 326–334

108. Van Lierde KM, Roeyers H, Boerjan S, De Groote I. Expressive and receptive language characteristics in three-year-old preterm children with extremely low birth weight. *Folia Phoniatr Logop*. 2009;61(5):296–299

109. Weindrich D, Jennen-Steinmetz C, Laucht M, Esser G, Schmidt MH. Epidemiology and prognosis of specific disorders of language and scholastic skills. *Eur Child Adolesc Psychiatry*. 2000;9(3):186–194

110. Whitehurst GJ, Arnold DS, Smith M, Fischel JE, Lonigan CJ, Valdez-Menchaca MC. Family history in developmental expressive language delay. *J Speech Hear Res*. 1991;34(5): 1150–1157

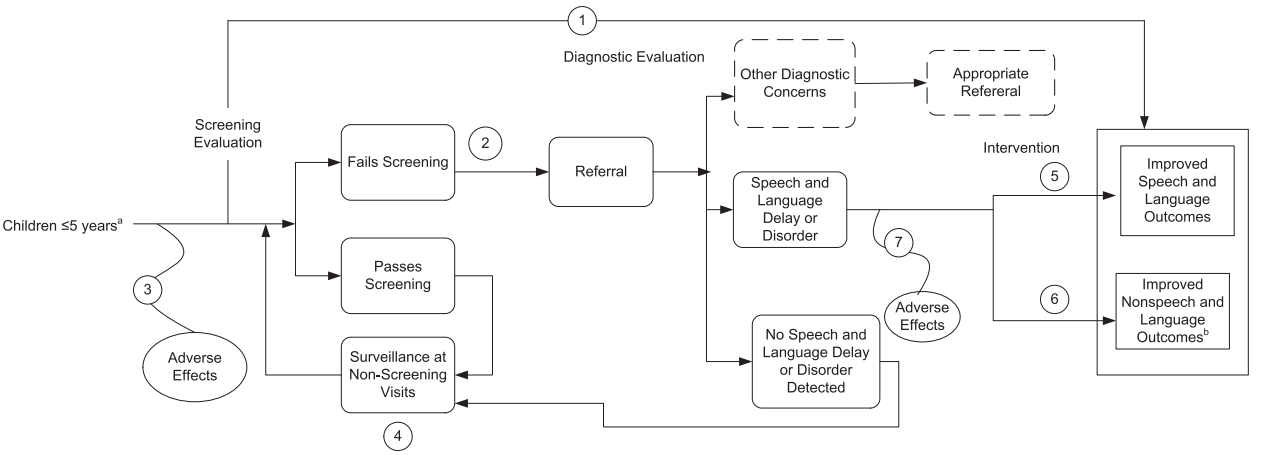
111. Yliherva A, Olsén P, Mäki-Torkko E, Koiranen M, Järvelin MR. Linguistic and motor abilities of low-birthweight children as assessed by parents and teachers at 8 years of age. *Acta Paediatr*. 2001;90(12): 1440–1449

112. Zambrana IM, Pons F, Eadie P, Ystrom E. Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *Int J Lang Commun Disord*. 2014; 49(3):304–316

113. Zubrick SR, Taylor CL, Rice ML, Slegers DW. Late language emergence at 24 months: an epidemiological study of prevalence, predictors, and covariates. *J Speech Lang Hear Res*. 2007;50(6): 1562–1592

Supplemental Information

Supplemental Figure 2. Analytic Framework and Key Questions



[a]Excluding children with diagnosed disorders including autism, mental retardation, Fragile X, hearing loss, degenerative and other neurologic conditions.
[b]School performance, behavioral competence, socioemotional development, quality of life, and others.

Key questions

1 Does screening for speech and language delay or disorders lead to improved speech and language outcomes as well as improved outcomes in domains other than speech and language?

2. Do screening evaluations in the primary care setting accurately identify children for diagnostic evaluations and interventions?

   a. What is the accuracy of these screening techniques and does it vary by age, cultural/linguistic background, whether the screening is conducted in a child's native language, or by how the screening is administered (i.e., parent report, parent interview, direct assessment of child by professional)?

   b. What are the optimal ages and frequency for screening?

   c. Is selective screening based on risk factors more effective than unselected, general population screening?

d. Does the accuracy of selective screening vary based on risk factors? Is the accuracy of screening different for children with an inherent language disorder compared with children whose language delay is due to environmental factors?

3. What are the adverse effects of screening for speech and language delay or disorders?
4. Does surveillance (active monitoring) by primary care clinicians play a role in accurately identifying children for diagnostic evaluations and interventions?
5. Do interventions for speech and language delay or disorders improve speech and language outcomes?
6. Do interventions for speech and language delay or disorders improve other outcomes such as academic achievement, behavioral competence, socioemotional development or health outcomes such as quality of life?
7. What are the adverse effects of interventions for speech and language delay or disorders (e.g., time, stress, and stigma)?

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Alberts et al., 1995[36] Fair 2006 review	Davis Observation Checklist for Texas Trained examiner	USA Head Start centers in central Texas	Children ages 52-67 months No description of recruitment methods Eligibility included normal hearing and English-language dominance No sampling for reference measure
Allen and Bliss, 1987[37] Fair 2006 review	Fluharty Preschool Screening Test Northwestern Syntax Screening Test Trained examiner	USA Child Care centers in suburban Detroit	Preschool-age children 36-47 months No description of recruitment methods No inclusion/exclusion criteria provided No sampling for reference measure
Bliss and Allen, 1984[38] Fair 2006 review	Screening Kit of Language Development Trained examiner	USA Child care centers in metropolitan Detroit	Preschool-age children 30-48 months No description of recruitment methods No inclusion/exclusion criteria provided No sampling for reference measure
Borowitz and Glascoe, 1986[39] Fair 2006 review	Denver Developmental Screening Test Trained examiner	USA Developmental evaluation center in middle Tennessee	Children ages 18-66 months Children were referred by Head Start centers, day care and preschool centers, public schools, public health agencies, the Department of Human Services, and private physicians Children were referred because of their home environment, medical problems, and suspected delays No sampling for reference measure
Burden et al., 1996[40] Good 2006 review	Parent Language Checklist Parent report	UK Community sample within Cambridge Health authority	Children aged 36 months Same sample as described by Stott et al., 2002[49], but it differs in terms of who was followed and analyzed. Families residing within the Cambridge Health Authority with a child born during a 9-month period were invited by mail to complete the screening when the child was 36 months. 1,936 of the 2,590 screeners were returned. Of the 472 Net-positive children, 277 proceeded to the screening. From the pool of randomly selected Net-negative children, 226 were randomly selected and 148 proceeded to the screening. A total of 425 children were included. Children were eligible if they were not a product of a multiple birth, had a listed medical condition (not described), lived in a multilingual home, or were eliminated on the telephone interview All children failing and not excluded and matched sample of children passing

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review (cont.)

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Drumwright et al., 1973[41] Fair 2006 review	Denver Articulation Screening Exam Trained examiner	USA Head Start, public and private child care centers, schools, pediatric clinics in Denver	Children ages 2.5-6 years No description of recruitment methods Children from economically disadvantaged homes No sampling for reference measure
Frisk et al., 2009[27] Fair New	Ages and Stages Questionnaire Communication domain Parent report Battelle Developmental Inventory Screening Test communication domain Brigance Preschool Screen Early Screening Profiles Trained examiner	Canada Child Development Programs (programs that provide early intervention services to young children at risk for developmental disabilities) in Ontario	Parents of children ages 4.5 years No description of recruitment methods Children were eligible if they were not legally blind, profoundly hearing impaired, untestable because of global delay, diagnosed with autism spectrum disorder, or used English as a second language with less than 19 months daily exposure to English Of the 131 children initially screened, data were available for only 111 children No sampling for reference measure
Guiberson and Rodriguez, 2010[28] Fair New	Spanish Ages and Stages Questionnaire Communication domain Pilot Inventario-III (Spanish Communicative Development Inventory) Parent report	USA Head Start centers, early childhood program, medical clinic in 2 western states	Parents of children ages 32-62 months Recruitment included sending flyers home to families with children enrolled in preschool programs, posting flyers in early childhood centers and medical clinics, and attending preschool family nights and Head Start (HS) community health fairs Eligible families spoke only or mostly Spanish; eligible children had normal hearing, no known neurological impairment, no severe phonological impairment, and spoke only or mostly Spanish Predetermined that approximately half of sample would have language delays and half without
Guiberson et al., 2011[29] Fair New	Spanish Ages and Stages Questionnaire Communication domain Short-form Inventarios del desarrollo de habilidades comunicativas: palabras y enunciados (Spanish CDI) Parent report	USA Early Head Start center, early intervention programs in 2 western states	Parents of toddlers ages 24-35 Study flyers sent to Early Head Start family members and service coordinators; interested parents of children in these programs who met inclusion criteria were invited Eligible families spoke only or mostly Spanish; eligible children had normal hearing, no known neurological impairment, and spoke only or mostly Spanish; children with both typical language development and expressive language delays were included Predetermined that approximately half of sample would have language delays and half without

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review (cont.)

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Heilmann et al., 2005[30] Study 2 Fair New	MacArthur-Bates Communicative Development Inventory : Words & Sentences Parent report	USA University research center	Parents of children who were 24 months Children were part of a larger longitudinal study of language delay who were recruited via birth registry, newspapers, flyers, posters at health fairs, and referrals from birth-to-3 providers Eligible children were from a monolingual English-speaking home, scored within the normal range on Denver II for general development, had normal hearing, and normal oral and speech motor abilities Included 38 late talkers and 62 children part of the larger study who had typical language
Klee et al., 1998[42] Fair 2006 review Klee et al., 2000[43] Fair	Language Development Survey Parent report	USA Community in Wyoming	Parents of children 24-26 months Families recruited by mail from 2 cities No inclusion/exclusion criteria provided All children who screened positive in an earlier study and a sample of those who screened negative were invited to participate in a comprehensive evaluation Same sample as in Klee et al., 1998[42], with a different analysis
Laing et al., 2002[44] Good 2006 review	Structured Screening Test Trained examiner	UK Health center in section of London	Children 30 months old Health visitors invited parents of all children who attended their child's 30-month developmental checkup to participate Children were eligible whether or not they had a previously diagnosed developmental disability No sampling for reference measure
Law, 1984[45] Good 2006 review	Hackney Early Language Screening Test Trained examiner	UK Pediatric practice in section of London	Children 30 months old All children attending a routine developmental checkup at age 30 months in a London suburb were screened No description of inclusion/exclusion criteria All who tested positive and a sample of those passed were seen for a diagnostic evaluation, provided their first language was English
Rescorla, 1989[46] Study 3 Fair 2006 review	Language Development Survey Parent report	USA University research center	Parents of children ages 23-34 months Parents recruited in response to a telephone inquiry following a notice in the paper and pediatricians' offices about a study of delayed language (delayed language sample) and through lists of participants in a previous study or whose pediatrician recommended them (typical language sample) Children recruited for study of language delay and the typical language comparison group

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review (cont.)

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Rescorla and Alley, 2001[47] Study 2 Fair 2006 review	Language Development Survey Parent report	USA University research center	Parents of children ages 23-34 months Sample of parents of children who were recruited for an epidemiological study of language delay in response a letter sent to all families of 2 year olds in 4 townships in a suburban Philadelphia county. The set who failed the LDS and a matched group who passed the LDS were invited to participate in Study 2. No inclusion/exclusion criteria described for epidemiological study other than age; for Study 2, sample of typical language children were matched to group with language delays on age, gender and SES All children who failed the LDS in the epidemiological study and a matched sample who passed the LDS
Rigby and Chesham, 1981[31] Fair New	A Trial Speech Screening Test Trained examiner	UK Primary care practice	4.5-year-old children Total population of children attending the school entrant medical examination Children were excluded if they were already receiving speech therapy No sampling for reference measure
Sachse and Von Suchodoletz, 2008[32] Sachse and Von Suchodoletz, 2009[33] Good New	ELFRA-2 (German version of MacArthur CDI: Toddler form) Parent report	Germany Community	Parents of 2-year-old children Parents recruited via birth announcements in a newspaper in Germany Children were eligible if they were from a monolingual German-speaking home and did not have poor vision, a hearing impairment, an abnormal result on a hearing screening, or missing subtests on the reference standard due to poor cooperation All children classified as late talkers based on the screening and a random sample of children with typical language development
Stokes, 1997[48] Good 2006 review	Developmental Nurse Screen Trained examiner Parent Questionnaire Parent	Australia Child Health Centers in metropolitan Perth	Parents of children 34- 40 months Letters were sent inviting parents along with a questionnaire Children were eligible if they had no developmental disability and English was their primary language Of the 1,500 parents invited, 409 consented, and 398 were included (11 were removed because of a developmental disability or non-English language) No sampling for reference measure

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review (cont.)

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Stott et al., 2002[49] Fair 2006 review	General Language Screen (formerly Parent Language Checklist) Parent report	UK Community within Cambridge Health authority	Children 36 months of age Families with a child born during a 9-month period were invited by mail to complete the screening 1,936 of the 2,590 screeners were returned, and 75 were excluded based on predefined (but unstated) criteria Selection of both passes and fails: 596 of 636 parents were interviewed at 37 months and 419 of the children were assessed at 39 months; 254 of 279 families who were invited were followed up at 45 months
Sturner et al., 1993[50] Fair 2006 review Study 1 Study 2	Fluharty Preschool Speech and Language Screening Test (Revision of Fluharty Preschool Screening Test) Trained examiner	USA School in a rural county in North Carolina School in a rural county in North Carolina	Children 53-68 months Parents recruited during kindergarten registration to bring their children back for screening; of the 378 who registered, 279 came for screening All kindergarten registrants Stratified samples of children completing the screening invited to return for testing - all positive and sample of borderline and negative screens Children 55-69 months Parents recruited during kindergarten registration to bring their children back for screening; of the 533 who registered, 421 came for screening All kindergarten registrants Stratified samples of children completing the screening invited to return for diagnostic testing - all positive screens and sample of borderline and negative screens
Sturner et al., 1996[51] Fair 2006 review	Sentence Repetition Screening Test Trained examiner	USA School in a rural county in North Carolina	Children 54-66 months Parents recruited during kindergarten registration to bring their children back for screening All kindergarten registrants Followup of all positive screens and sample of borderline and negative screens
Ward, 1984[52] Fair 2006 review	Ward screening tool (author-created) Trained examiner	UK Community in one district in Manchester	Children 7-23 months old All parents in district were invited to a local clinic for a hearing test between the ages of 7 and 9 months (screening occurred between 7 and 23 months). Children were ineligible if their caregivers had limited English No sampling for reference measure

Supplemental Table 3. Screening Accuracy Studies from 2006 Review and New Review (cont.)

Study Reference Quality Rating Source	Screening Tool Screening Source	Country, Recruitment Setting	Sample Description, Recruitment Method, Inclusion/Exclusion Criteria, Sampling for Reference Measure
Westerlund et al., 2006[34] Fair 2006 review	Swedish Communication Screening -18 (derived from Swedish MacArthur-Bates CDI) Parent report	Sweden Community sample invited to all child health centers in one county	Parents of children 18 months old All parents of 18-month-old children invited to come to child health care centers based on the national population register of the region All had Swedish as their primary language No sampling for reference measure
Wetherby et al., 2003[35] Fair New	Infant-Toddler Checklist from the Communication and Symbolic Behavior Scales Parent report	USA Research sample recruited from the community for a longitudinal study	Parents of children 6 - 24 months old Parents recruited from public announcements, community family events, health care providers, child care providers, public agency that provides services infants and toddlers under Part C of the Individuals with Disabilities Education Act Sample was drawn from 2434 parents who completed the Infant-Toddler Checklist and the subset of 392 children 12-24 mos. whose parents also completed a Behavior Sample Inclusion criteria included completion of Behavior Sample within 2 mos. of the Infant Toddler Checklist All children who failed the screen and samples of those scored between the mean and 1 SD below the mean and those who scored at or above the mean

LDS, Language Development Survey; SD, standard deviation; UK, United Kingdom; USA, United States of America.

Supplemental Table 4. Screening Tools for Speech and Language Delay and Disorders for Children 5 years and younger

Screening Tool	Domain(s) or Skills Screened	Summary Scores	Number of Items Time to Complete	Appropriate Ages	Reliability	Screening Source
Ages and Stages Questionnaire Communication domain, 2nd edition[27,72]	Broad communication skills	• Communication	6 (at each of 19 age levels) NR	4-60 months	Coefficient alpha = 0.63-0.75	Parent report
Spanish version of Ages and Stages Questionnaire Communication domain[28]	Broad communication skills in Spanish	• Communication	6 (at each of 19 age levels) NR	4-60 months	NR	Parent report
Battelle Developmental Inventory Screening Test Communication domain[27]	Receptive and expressive language skills	• Receptive • Expressive	9 9 NR	12-96 months	NR	Trained examiner
Brigance Preschool Screen[27]	Receptive and expressive language skills	• Understanding reading (receptive language) • Expressive language	2 4 NR	45-56 months	NR	Trained examiner
Davis Observation Checklist for Texas[36]	Speaking, understanding. speech fluency, voice and hearing	• Communication	2-5 behaviors (in each of 6 areas) NR	4-5 years	NR	Trained examiner
Denver Articulation Screening Exam[41]	Articulation	• Articulation	34 sound elements NR	2.5-7 years	Test-retest = 0.95	Trained examiner
Denver Developmental Screening Test Language Sector[39]	Broad language skills	• Global language	NR NR	1 month - 6 years	NR	Trained examiner
Developmental Nurse Screen[48]	Broad language	• Global language	1 NR	34-40 months	NR	Trained examiner
Early Screening Profiles[27]	Word comprehension and production	• Verbal concepts	25 NR	2-0 - 6-11 years	NR	Trained examiner
Fluharty Preschool Screening Test/ Fluharty Preschool Speech and Language Screening Test[50,73]	Articulation, expressive, and receptive language skills	• Articulation • Language	35 6-10 minutes	2-6 years	Test-retest = 0.96 - 0.98	Trained examiner

Supplemental Table 4. Screening Tools for Speech and Language Delay and Disorders for Children 5 years and younger (cont.)

Screening Tool	Domain(s) or Skills Screened	• Summary Scores	Number of Items Time to Complete	Appropriate Ages	Reliability	Screening Source
General Language Screen (GLS)/Parent Language Checklist[40,49]	Comprehension, expression, articulation, pragmatics	• Global language	11 NR	36 months	Coefficient alpha = 0.74	Parent report
Infant-Toddler Checklist[35]	Emotion and eye gaze, communication, gestures, sounds, words, understanding, object use	• Social composite • Speech composite • Symbolic composite • Total score	24 5-10 minutes (entire screener)	6-24 months	Test-retest Total = 0.86 Social = 0.70 Speech = 0.73 Symbolic = 0.79	Parent report
Language Development Survey[46,47]	Expressive vocabulary and word combinations	• Expressive language	310 words, word combinations NR	18-35 months	Coefficient alpha = 0.99 Test-retest = 0.86 to 0.99 for categories	Parent report
MacArthur-Bates Communicative Development Inventory (CDI): Words and Sentences(W-S)[30,74]	Expressive vocabulary, morphology, and grammar	• Productive vocabulary	798 words, morphological markers, Sentences 20–40 minutes	16-30 months	Test-retest Complexity = 0.96 Vocabulary = 0.95	Parent report
ELFRA-2 German version of MacArthur CDI: Toddler (now CDI: Words and Sentences)[32]	German expressive vocabulary, morphology, and grammar	• Global language (using all components)	260 vocabulary 25 syntax 11 morphology NR	16-30 months	NR	Parent report
Pilot Inventario-III (Spanish version of CDI-III)[28]	Expressive vocabulary, grammar, usage	• Expressive language	100 vocabulary 12 sentence usage 12 language use NR	30-37 months	Coefficient alpha Vocabulary = 0.92 Sentences = 0.95 Usage = 0.94	Parent report
Short Form of Inventarios del desarrolo de habilidades comunicativas: palabras y enuciados: Spanish version of CDI - WS[29,75]	Spanish expressive vocabulary, morphology, and grammar	• Expressive language	100 words, word combinations 15 minutes	16-30 months	NR	Parent report
Swedish Communication Screening (SCS-18) (derived from Swedish CDI)[34]	Swedish expressive and receptive vocabulary, morphology, grammar	• Word production	90 words 13 gestures NR	18 months	Coefficient alpha Word production= 0.97 Word comprehension = 0.96 Test–retest Word production = 0.97 Word comprehension = 0.89	Parent report

Supplemental Table 4. Screening Tools for Speech and Language Delay and Disorders for Children 5 years and younger (cont.)

Screening Tool	Domain(s) or Skills Screened	Summary Scores	Number of Items Time to Complete	Appropriate Ages	Reliability	Screening Source
Northwestern Syntax Screening Test[37]	Expressive and receptive knowledge of syntactic forms	Syntactic expression Syntactic comprehension	40-20 expressive and 20 receptive NR	3-8 years	NR	Trained examiner
Parent Questionnaire[48]	Sentence use Comprehension Articulation Problems	Global language	4 2 minutes	34-40 months	NR	Parent report
Screening Kit of Language Development[38]	Vocabulary comprehension, story completion, sentence completion, paired sentence repetition, individual sentence repetition with and without pictures, comprehension of commands	Global language	38-50 items per subtest 10 minutes	30-60 months	NR	Trained examiner
Sentence Repetition Screening Test[51,76]	Expressive morphology and articulation	Global language Articulation	15 NR	54-66 months	Coefficient alpha Language = 0.83 Articulation = 0.88	Trained examiner
Trial Speech Screening Test[31]	Articulation Grammar	Language	12 NR	54 months	NR	Trained examiner
Ward screening tool(author-created)[52]	Attention to auditory and language stimuli, prelanguage expression	Prelinguistic behaviors	10 NR	7-9 months	NR	Parent report

NR, not reported.

Supplemental Table 5. Characteristics of randomized controlled trials of speech and language interventions

Study, Country, Risk of Bias	Speech and Language Domains	Intervention	Length of Intervention, Timing of Outcome Assessment	Inclusion/ Exclusion Criteria	Child's Age at Baseline in Months	N Patients Randomized
Almost et al., 1998[66] Canada Fair	Speech sounds (phonology) Language (expressive)	G1: Clinician-directed individualized therapy G2: Delayed tx	Two 30-minute sessions per week for 4 mos (1,040 mins total). Outcome assessment at 4 mos.	Inclusion: severe phonological disorder, normal receptive language, hearing, oral structures and function, and sufficient attention span	G1: 42.5 (Range: 33-61) G2: 42.5 (Range: 33-55)	Overall: 26 G1: 13 G2: 13
Fricke et al, 2013[54] United Kingdom Fair	Language (expressive and receptive) Speech sounds (phonological awareness)	G1: Oral language group intervention to teach skills related to vocabulary, expressive vocabulary and grammatical competence, to encourage active listening and build confidence in independent speaking G2: Usual nursery/primary school activities	3 15-min group sessions per wk for 10 wks, increasing to 5 sessions per wk (3 30-min group sessions plus 2 15-min individual sessions) for 20 wks (2,850 mins total). Outcome assessment at end of tx (30 wks) and at 6 mo follow-up	Inclusion: Twelve children in each of 15 nursery schools with the lowest mean verbal composite scores	G1: 48 G2: 48 (Screening occurred at 48 months)	Overall: 180 G1: 90 G2: 90
Gibbard, 1994[61] United Kingdom Fair	Language (expressive and receptive)	G1:Parent group S&L training, mix of approaches focusing on activities for parent to use with children, many from the Derbyshire Language Scheme G2: Wait list	60-75 mins every other wk for 6 mos (780-975 mins total). Outcome assessment at 6 mos.	Inclusion: Age 27-39 mo, little or no expressive language, no: general developmental delay, medical condition indicative of a language delay, previous S&L therapy	G1: 35 (Range: 29-39) G2: 32 (Range: 27-39)	Overall: 36 G1: 18 G2: 18
Girolametto et al., 1996[62] Companion: Girolametto et al., 1997[67] Canada Fair	Language (expressive); Speech sounds (phonology)	G1: Hanen Program for Parents training program modified to be consistent with a focused stimulation of children's language G2: Wait list	Eight 150 min parent group sessions (1,200 mins total) and 3 home visits over 11 wks. Outcome assessment 3 wks following end of tx	Inclusion: Expressive language delay, single-word stage of language development, only English spoken in home Exclusion: major sensory impairment, oral motor problems, neurological problems, ASD	G1: 28.7 (Range:25-35) G2: 28.6 (Range 23-34)	Overall: 25 G1: 12 G2: 13

Supplemental Table 5. Characteristics of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Intervention	Length of Intervention, Timing of Outcome Assessment	Inclusion/ Exclusion Criteria	Child's Age at Baseline in Months	N Patients Randomized
Glogowska et al., 2000[65] United Kingdom Good	Language (expressive and receptive), Speech sounds (phonology)	G1: Individually tailored "routine" S&L therapy by a therapist G2: Wait list "watchful waiting"	Avg of 6.2 hr of therapy over 12 mo (372 min total). Outcome assessment at 12 mo	Inclusion: Preschoolers in S&L therapy based on general or expressive language group scores on preschool language scale or phonology group scores	G1: 34.2 (Range: 18-42) G2:34.2 (Range: 24-42)	Overall: 159 G1: 71 G2: 88 (18 crossed over before study end)
Jones et al., 2005[55] New Zealand Fair	Fluency	G1: Lidcombe Program of Early Stuttering according to the manual G2: Delayed tx	Conducted in 2 stages. During stage 1, parent conducted program each day and speech pathologist once per wk. Outcome assessment at 9 mo	Inclusion: Age 3-6 years, diagnosed stuttering with 2+% of syllables stuttered, English-speaking Exclusion: stuttering: tx in previous 12 mo, onset in 6 mo before recruitment	G1: 56.4 G2: 46.8 (Range: 36-72)	Overall: 54 G1: 29 G2: 25
Lewis et al., 2008[56] Australia Fair	Fluency	G1: Lidcombe Program of Early Stuttering, a manualized intervention delivered through telehealth (phone, video and audio recordings) G2: Delayed tx	Typically at least 1 wkly phone consultation; video demonstrations, phone and mail support. Outcome assessment at 9 mos	Inclusion: stuttering for longer than 6 mo, no current or previous tx, all other development normal, parent and child English-speaking	Means: NR (Range: 36-54)	Overall: 22 G1: 9 G2: 13
Morgan and Vogel, 2009[60] Australia Systematic Review Fair	Childhood apraxia of speech	RCT tx studies of interventions delivered by S&L therapists	No studies met inclusion criteria	Include: Age 3-16 yrs	NA	NA

Supplemental Table 5. Characteristics of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Intervention	Length of Intervention, Timing of Outcome Assessment	Inclusion/ Exclusion Criteria	Child's Age at Baseline in Months	N Patients Randomized
Robertson et al., 1997[64] United States Fair	Language (expressive and receptive)	G1: Unstructured play sessions in "house" area with normal peers G2: No play sessions with normal peers	At least 4 15-20 min session over 3 wk (minimum of 60 minutes). Outcome assessment: 3 wk following end of tx	Inclusion: Language impairment and in language-based early childhood classroom, WISC-R score: 85+, poor receptive and expressive language, no motor, emotional or physical handicaps, no hearing or vision problems, monolingual English	G1: 49.8 G2: 49.6 (Overall range: 44-61)	Overall: 20 G1: 10 G2: 10
Robertson et al., 1999[68] United States Fair	Language (expressive and receptive) Speech sounds (intelligibility)	G1:Speech-language pathologist directed small group therapy of no more than 4 children G2: Wait list	150 mins per wk for 12 wks (1,800 mins total). Outcome assessment at end of tx.	Include: normal hearing, oral and speech motor abilities, no frank neurological impairment, monolingual, English-speaking homes, CDI vocabulary scores < 10th percentile, no other delays, no prior S&L therapy	G1: 25.6 (Range: 21-30) G2: 24.6 (Range: 21-28)	Overall: 21 G1: 11 G2: 10
Shelton et al., 1978[63] United States Fair	Speech sounds (phonology, articulation) Language (expressive and receptive)	G1: Parent-directed speech sound listening/discrimination activities (listening group) G2: Parent-child storybook interaction (reading and talking group) G3: Control group	G1: 5 min per day, 5 days per wk for 57 days (1,425 mins total) G2: 15 min per day, 5 days per wk for 57 days (4,275 ms total). Outcomes assessment: end of tx	Inclusion: Below cut-off score on Templin-Darley Articulation Screening Test, pass audiometric screening	G1: 47 G2: 49 G3: 39 (Overall range: 27-55)	Overall: 45 G1: 15 G2: 15 G3: 15

Supplemental Table 5. Characteristics of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Intervention	Length of Intervention, Timing of Outcome Assessment	Inclusion/ Exclusion Criteria	Child's Age at Baseline in Months	N Patients Randomized
Wake et al., 2011[57] Clustered Randomized Trial Australia Good	Language (expressive and receptive)	G1: Modified "You Make the Difference" (Hanen parent training program): Low intensity version of parent-delivered toddler language promotion program for toddlers identified as slow to talk on universal screening. G2: Usual care (not defined)	120 min per wk for 6 wks (720 mins total). Outcome assessment: child age 2 yrs (12-14 wks following program completion) and child age 3 yrs	Inclusion: at or below 20th percentile in expressive vocabulary at 18 mo. Exclusion: Cognitive delay, major medical conditions, or suspected ASD; parents with insufficient English.	G1: 18.1 (SD=0.7) G2: 18.1 (SD=0.8)	Overall: 301 G1: 158 G2: 143
Wake et al., 2013[59] Australia Fair	Language (expressive and receptive), Speech sounds (phonological awareness)	G1: 18 1-hour home-based therapy sessions conducted by a "language assistant" G2: No intervention control; "free to participate in community-based txs	Eighteen in-home 1-hr targeted sessions in 3 blocks of wkly sessions for 6 wks starting every 3 mo	Inclusion: expressive and/or receptive language scores >1.25 SD below normal, Exclusion: intellectual disability, major medical conditions, hearing loss >40 dB in the better ear, ASD, parents with insufficient English	G1: 50 G2: 49 (recruited at 48 mo)	Overall: 200 G1: 99 G2: 101
Yoder et al., 2005[58] United States Fair	Language (expressive), Speech sounds (intelligibility)	G1: Broad target recasts intervention G2: No intervention control; "free to participate in community-based txs"	Three 30-min tx sessions per wk for 6 mo (2,340 min total). Outcome assessment: at end of tx and at followup (8 mo later)	Inclusion: Specific language and speech accuracy impairments; nonverbal IQ > 80; no hearing impairment; monolingual English, no oral motor disorders	G1: 44.3 (SD=7.6) G2: 43.2 (SD=9.6)	Overall: 52 G1: 26 G2: 26

ASD = autism spectrum disorder; avg = average; CDI = MacArthur-Bates Communicative Development Inventories; dB = decibels; G = group; min = minute; mo = month; N = number; NR = not reported; RCT = randomized controlled trial; SD = standard deviation; tx = treatment; WISC-R = Wechsler Intelligence Scale for Children; wk = week; yr = year

Supplemental Table 6. Outcomes of randomized controlled trials of speech and language interventions

Study, Country, Risk of Bias	Speech and Language Domains	Outcomes Speech and Language (KQ5) Non-Speech and Language (KQ6)	Summary of Findings
Almost et al., 1998[66] Canada Fair	Speech sounds (phonology) Language (expressive)	Diff measured through ANCOVA to adj for baseline, tx and time[a] **Speech and Language (KQ5)** Phonological processes (APP-R): F=8.64, d=1.15 (p=0.007) Articulation (GFTA): F=8.92, d=1.17 (p=0.007) Consonants correct (PCC): F=8.06, d=1.11 (p=0.009) MLU: F=0.23, d=.18 (p=0.638) **Non-Speech and Language (KQ6)** None reported	S&L: More improvement in 3 of 4 measures of phonology Non-S&L: No measures reported
Fricke et al, 2013[54] United Kingdom Fair	Language (expressive and receptive) Speech sounds (phoneme awareness)	Diff measured through structural equation modeling to allow for missing data and clustering of children within schools **Speech and Language (KQ5)** Language: end of tx: d=0.80 (p<0.01); follow-up: d=0.83 (p<0.001) Narrative: end of tx: d=0.39 (p=0.003); follow-up: d=0.30 (p=0.041) Phoneme awareness: end of tx: d=0.49 (p<0.031); follow-up: d=0.49 (p=0.01) **Non-Speech and Language (KQ6)** Literacy: end of tx: d=0.31 (p=0.07); follow-up: d=0.14 (p=0.354) Letter knowledge: end of tx: d=0.41 (p<0.001) Diff in reading comprehension at follow-up: 0.97 (95% CI, 0.40 to 1.54), d=0.52, (p=0.001)	S&L: Better performance on language, narrative, and phoneme awareness at posttest and at 6-mo follow-up. Non-S&L: Better reading comprehension but no diff in reading accuracy at 6 mo follow-up
Gibbard, 1994; Study 1[61] United Kingdom Fair	Language (expressive and receptive)	All diff measured through ANCOVA to adj for baseline, tx, and time[a] **Speech and Language (KQ5)** Reynell Expressive: F = 64.89, d=2.69 (p<0.001) Reynell Comprehension: F=34.11, d=1.95 (p<0.001) Derbyshire One Word Scores: F=34.24, d=1.95 (p<0.001) Derbyshire Total Scores: F=31.94, d=1.88 (p<0.001) Renfrew Grammatical Ability: F=20.36, d=1.50 (p<0.001) Renfrew Information: F=32.0, d=1.89 (p<0.001) MLU: F=24.44, d=1.65 (p<0.001) **Non-Speech and Language (KQ6)** None reported	S&L: More improved in all measures of S&L Non-S&L: No measures
Girolametto et al., 1996[62] Companion: Girolametto et al., 1997[67] Canada Fair	Language (expressive); Speech sounds (phonology)	Diff measured through MANCOVA to adj for baseline, tx and time[a] **Speech and Language (KQ5)** Expressive vocabulary: Size: F=4.90, d=0.88 (p<0.01) Number of diff words in interaction: F=7.96, d=1.13 (p<0.02) Number of learned control words: F=17.25, d=1.67 (p<0.01) Talkativeness: F=2.38, d=0.62 (p<0.06) Parent report on structural complexity: F=2.85, d=0.68 (p<0.04) Consonant inventory: F=4.34 (p<0.01) Early consonants, d=1.0; middle consonants, d=1.1; late consonants, d= 0.6 Percent of consonants correct: d = -0.3 (p = NS) Number of vocalizations:(p =NS) Syllable structure level (Level 3 vocalizations): F=6.74, d=.9 (p<0.01)	S&L: More improved in measures of vocab size, use of more diff words, more structurally complete and complex utterances, more multiword utterances, and larger inventory of consonants. No diff in number of vocalizations or rate of words per min

Supplemental Table 6. Outcomes of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Outcomes Speech and Language (KQ5) Non-Speech and Language (KQ6)	Summary of Findings
Girolametto et al., 1996[62] Companion: Girolametto et al., 1997[67] Canada Fair (continued)		**Non-Speech and Language (KQ6)** None reported	Non-S&L: No measures
Glogowska et al., 2000[65] United Kingdom Good	Language (expressive and receptive), Speech sounds (phonology)	All diff measured through ANCOVA to adj for baseline, tx and time (95% CI)[a] **Speech and Language (KQ5)** Auditory comprehension, avg of diff at 6 & 12 mo: 4.1 (0.5 to 7.6) d=~0.3 (p=0.025) Expressive language, avg of diff at 6 & 12 mo: 1.4 (-2.1 to 4.8) (p =0.44) Phonology error rate, avg of diff at 6 & 12 mo: -4.4 (-12.0 to 3.3) (p=0.26) Bristol language development scale, avg of diff at 6 & 12 mo: 0.1 (-0.4 to 0.6) (p=0.73) Improvement by 12 mo on clinical criteria used for study entry: OR =1.3 (0.67 to 2.4) (p=0.46) **Non-Speech and Language (KQ6)** Wellbeing, avg of diff at 6 & 12 mo: 0.04 (-0.2 to 0.3) Attention level, avg of diff at 6 & 12 mo: 0.02 (-0.3 to 0.3) (p=0.91) Play level, avg of diff at 6 & 12 mo: 0.04 (-0.2 to 0.2) Vineland Socialization Scale, avg of diff at 6 & 12 mo: 0.6 (-3.1 to 4.2)	S&L: More improved auditory comprehension, no diff in expressive language, phonology error rate, language development Non-S&L: No diff in well-being, attention, play level or socialization skills
Jones et al., 2005[55] New Zealand Fair	Fluency	Percent syllables stuttered: adj mean diff (95% CI)[a] **Speech and Language (KQ5)** 2.3 (0.8 to 3.9) (p=0.003) Diff measured through logistic regression to adj for baseline Odds of<1% of syllables stuttered: OR=0.13 (0.03 to 0.63) (p=0.011) **Non-Speech and Language (KQ6)** None reported	S&L: Greater reduction in % of syllables stuttered and greater odds of stuttering less than 1% of syllables. Non-S&L: No measures
Lewis et al., 2008[56] Australia Fair	Fluency	Diff measured through ANCOVA to adjust for baseline, tx, and time (95% CI) **Speech and Language (KQ5)** Stuttering frequency: At 9 mo: 69% (13% to 89%) (p=0.04) Adjusting for patient characteristics: 73% (25% to 90%) (p=0.02) **Non-Speech and Language (KQ6)** None reported	S&L: Greater reduction in % of syllables stuttered during speech sample; more "responders" (i.e., decrease of >80% in stuttered syllables) Non-S&L: No measures

Supplemental Table 6. Outcomes of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Outcomes Speech and Language (KQ5) Non-Speech and Language (KQ6)	Summary of Findings
Robertson et al., 1999[68] United States Fair	Language (expressive and receptive) Speech sounds (intelligibility)	Diff measured through ANCOVA to adjust for baseline, tx, and time[a] **Speech and Language (KQ5)** MLU: F=10.33, d=1.40 (p<0.003) Total number of words: F=46.83, d=2.99 (p<0.001) Number of diff words: F=41.05, d=2.80 (p<0.001) Number of diff words, controlling for number of words: F=24.03, d=2.14 (p<0.001) Lexical repertoire: F=46.86, d=2.99 (p<0.001) Percentage of intelligible utterances: F=24.44, d=2.16 (p<0.001) **Non-Speech and Language (KQ6)** Vineland Socialization Scale: F=12.15, d=1.52 (p=0.003) Parental stress (Child domain of the PSI): F=53.32, d=3.19 (p<0.001)	S&L: More improvement in MLU, number of words, vocabulary size, lexical diversity, and % of intelligible utterances Non-S&L: More improvement in socialization skills, greater reduction in parental stress
Robertson et al., 1997[64] United States Fair	Language (expressive and receptive)	Diff measured through ANCOVA to adjust for baseline, tx, and time[a] **Speech and Language (KQ5)** Number of words: F=70.72 (p<0.01) Number of different words: F=73.79 (p<0.01) Play-theme-related acts: F=99.80 (p<0.01) Linguistic markers: F=73.51 (p<0.01) **Non-Speech and Language (KQ6)** None reported	S&L: More words used, greater verbal productivity, more lexical diversity, reported play acts, and linguistic markers Non-S&L: No measures
Shelton et al., 1978[63] United States Fair	Speech sounds (phonology, articulation) Language (expressive and receptive)	Diff between the 3 groups measured through ANCOVA to adjust for baseline, tx, and time[a], Results always in comparison to control **Speech and Language (KQ5)** Test of Auditory Discrimination (quiet): Listening: d= 0.17; Reading & Talking: d=-0.05 (p=0.90) Test of Auditory Discrimination (noise): Listening: d=-0.41; Reading & Talking: d= 0.91; (p=0.03) (greatest improvement in controls) Northwestern Syntax Screening Test: Listening: d= -0.17; Reading & Talking: d= 0.10 (p=0.72) Auditory Association Subtest of ITPA: Listening: d= 0.50; Reading & Talking: d= 0.51 (p =0.25) Discrimination Task: Listening: d=0; Reading & Talking: d=-0.05 (p=1.00) Error Recognition: Listening: d= 0.17; Reading & Talking: d=0.40 (p =0.26) Templin-Darley Articulation Screening Test: Listening: d= 0.65; Reading & Talking: d=.02 (p=0.07) McDonald Screening Deep Test of Articulation: Listening: d= 0.06 Reading & Talking: d= 0.-38 (p=0.51) **Non-Speech and Language (KQ6)** None reported	S&L: No diff in improvements between intervention groups and controls in relation to articulation, auditory discrimination, or auditory association Non-S&L: No measures

Supplemental Table 6. Outcomes of randomized controlled trials of speech and language interventions (cont.)

Study, Country, Risk of Bias	Speech and Language Domains	Outcomes Speech and Language (KQ5) Non-Speech and Language (KQ6)	Summary of Findings
Wake et al., 2011[57] Australia Good	Language (expressive and receptive)	All diff measured through random effects regression to adj for clustering, potential confounders, baseline measures (95% CI)[a] **Speech and Language (KQ5)** MCDI vocabulary raw score At 2 yr: 2.1 (−3.0 to 7.2), d=0.004 (p=0.42) At 3 yr: 4.1 (-2.3 to 10.6), d=0.08 (p=0.21) PLS expressive communication standard score: At 2 yr: 1.2 (−1.6 to 4.0), d=0.02 (p=0.41) EVT expressive vocabulary standard score: At 3 yr: −0.5 (−4.4 to 3.4), d=-0.08 (p=0.80) PLS auditory comprehension standard score At 2 yr: 1.4 (−2.2 to 5.0), d=-0.01 (p=0.44) At 3 yr: −0.3 (−4.2 to 3.7), d=-0.06 (p=0.90) **Non-Speech and Language (KQ6)** CBCL externalizing behavior raw score At 2 yr: −0.3 (−1.6 to 1.1), d=-0.04, p=0.71) At 3 yr: −0.1 (−1.6 to 1.4), d=-0.01,(p=0.86) CBCL internalizing behavior raw score At 2 yr: 0.1 (−0.9 to 1.1), d=-0.06,(p=0.78) At 3 yr: -0.1 (−1.3 to 1.2), d=-0.06, (p=0.92)	S&L: No diff in expressive or receptive language outcomes at 2 or 3 yr of age Non-S&L: No diff in internalizing or externalizing problem behaviors at 2 or 3 yr of age
Wake et al., 2013[59] Australia Fair	Language (expressive and receptive), Speech sounds (phonological awareness)	Mean diff (95% CI) measured at age 5, adjusting for gender, mother's education, recruitment from Let's Read or Let's Learn Language, baseline expressive and receptive language scores, and baseline measure of the outcome being considered, when available. **Speech and Language (KQ5)** Expressive language: 2.0 (-0.5 to 4.4), d=0.2, (p=0.12) Receptive language: 0.6 (-2.5 to 3.8), d=0.05, (p=0.69) Pragmatic language: -1.0 (-3.7 to 1.6), d=-0.1, (p=0.45) Phonological awareness: 5.0 (2.2 to 7.8), d=0.6, (p<0.001) **Non-Speech and Language (KQ6)** Letter knowledge: 2.4 (0.3 to 4.5), d=0.3, (p=0.03) Number of behavior problems: -0.5 (-1.7 to 0.7), d=-0.1, (p = 0.43) Health-related quality of life: -0.8 (-5.2 to 3.5), d=-0.05, (p=0.71)	S&L: No diff in expressive or receptive language outcomes. Better phonological awareness Non-S&L: Better letter knowledge. No diff in behavior problems or health-related quality of life
Yoder et al., 2005[58] US Fair	Language (expressive), Speech sounds (intelligibility)	Diff measured through ANCOVA to adj for baseline, tx and time **Speech and Language (KQ5)** At end of tx: (p=NS) Eight mo following end of tx: (p=NS) Among children who began tx with lowest articulation scores, diff in MLU at end of tx (p=0.01) and at 8 mo follow-up (p=0.03) **Non-Speech and Language (KQ6)** None reported	S&L: No diff between groups in change over time Non-S&L: No measures

[a]Cohen's d calculated by the review authors.

Adj = adjusted; ANOVA = analysis of variance; APP-R: Assessment of Phonological Processes-Revised; CBCL = Child Behavior Checklist; CI = confidence interval; d = Cohen's d; diff = difference; EVT = Expressive Vocabulary Test; GFTA = Goldman-Fristoe Test of Articulation; ITPA = Illinois Test of Psycholinguistic Ability; MCDI= MacArthur-Bates Communicative Development Inventory; MLU = mean length of utterances; MANCOVA = multivariate analysis of variance; mo, months; NA, not applicable; NS = not significant; OR = odds ratio; PCC = percentage consonants correct; PLS = Preschool LLanguage Scale; PSI = Parental Stress Index; S&L: speech and language; tx, treatment; VABS = Vineland Adaptive Behavior Scales; yr, years.

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education

Author, Year Analysis Approach[a]	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Adams-Chapman et al., 2013[77] Multivariate	Language delay composite measure (expressive and receptive)	Preterm infants born ≤26 weeks included in US Neonatal Research Network Follow-Up Study (N = 1,477)	18-22 months	NR	NR	↑	NR	NR	Mom < high school: ↑
Alston and St. James-Roberts, 2005[78] Univariate	Language and communication difficulty	Infants who completed the WILSTAAR early language and communication screening assessment (N=60)	Mean age: Not at-risk: 9.4 months, At-risk: 10.0 months	NR	NR	0	0	Dad age: ↑	NR
Campbell, 2003[79] Multivariate	Speech delay	Cohort of children being followed to study otitis media, Pittsburgh (N=639)	36 months	NR	↑	↑	NR	NR	Low Mom: ↑
Choudhury and Benasich, 2003[80] Case control: univariate	Low language as measured by PLS-3: Expressive, receptive and total score; Stanford-Binet: Verbal vocabulary, verbal comprehension; CELF-P: Word structure, sentence structure	Cohort with family history of specific language impairment and matched controls, New York City area (N=92)	36 months	NR	All measures except CELF-P sentence structure:↑	↑	NR	NR	NR
Desmarais et al., 2008[81] Analysis approach varies by study	Late talking	Review of 25 publications	18-39 months	NR	↑	NR	↑	NR	0

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Everitt, Hannaford and Conti-Ramsden, 2013[82] Case control	Persistent expressive language delay vs. typical language development among children with specific expressive language delay 1year earlier	Nursery school children in Scotland (N=94)	4-5 years	specific expressive language delay: 0 Received S&L therapy: 0 Poorer performance on PLS-3 AC, PLS-3 EC and Recalling Sentences subtest 1 year earlier	0	0	NR	NR	Mother: 0 Father: 0
Foster-Cohen, et al., 2010[83] Multivariate	Poorer receptive and expressive language ability	Very preterm cohort, compared to full-term born in New Zealand (N=204)	4 years	NR	NR	NR	NR	NR	NR
Fox, Dodd, and Howard, 2002[84] Multivariate	Functional speech disorders	German cohort (N=113)	32-86 months	NR	↑	NR	NR	NR	NR
Glascoe and Leew, 2010[85] Multivariate	Delay in Communication (Expressive and receptive language)	US nationally representative sample included in Brigance Infant and Toddler Screens study (N=382)	2 weeks-24 months	NR	NR	NR	Employ-ment: 0	NR	0
Hammer, 2010[86] Multivariate	Parent-reported speech-language impairment	Head Start Family and Child Experiences Survey cohort (N=1,015)	3-4 years (mean: 50 months)	NR	NR	↑	NR	NR	0

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Harrison and McLeod, 2010[87] Multivariate	Expressive speech and language concern, receptive speech and language concern, low receptive vocabulary	Growing Up in Australia- The Longitudinal Study of Australian Children cohort (N=4,980)	51-67 months, 80% were 54-60 months	NR	NR	All 3 outcomes: ↑	Household income: Expressive and receptive: 0 : Vocabulary : ↓ Financial hardship: all outcomes: 0	Expressive and receptive: 0 Vocabulary : ↓	Mom: Expressive and receptive: 0, Vocabulary: ↓ Dad: All outcomes: 0
Henrichs, et al., 2011[88] Multivariate	Expressive vocabulary delay (late bloomers, late onset, or persistent delay)	Generation R Study cohort, Netherlands (N=3,759)	Mean: 31.6 months	Receptive delay at 18 months: ↑	NR	NR	0	Late bloomer:↑ Late onset: ↓	Late onset:↓ Persistent: ↓
Kerstjens, et al., 2011[89] Multivariate	Ages and Stages Communication domain delays	Community-based and preterm cohorts in the Netherlands (N=1,983)	43-49 months (4 year old assessment)	NR	NR	Included in model but NR	NR	NR	In model but NR
Kerstjens, et al., 2009[90] Multivariate	Ages and Stages Communication domain delays	Community-based and preterm cohorts in the Netherlands (N=1,893)	43-49 months (4 year old assessment)	NR	NR	↑	Low: ↑	NR	0
Kerstjens, et al., 2012[91] Multivariate	Ages and Stages Communication domain delays	Community-based and preterm cohorts in the Netherlands (N=1,983)	43-49 months (4 year old assessment)	NR	NR	Included in model but NR	NR	NR	In model but NR
Law, et al., 2009[15] Multivariate	Specific language impairment (SLI); nonspecific language impairment (N-SLI)	British Cohort Study cohort (N = 9,132)	5 years (60 months)	Ever seen speech and language therapist: SLI ↑, N-SLI ↑	NR	SLI: ↓, N-SLI: ↓	Overcrowding: SLI: ↑, N-SLI: ↑	NR	Mom: SLI: 0; N-SLI: ↓ Parent poor reader: SLI: ↑, N-SLI: ↑

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year Analysis Approach	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Law, et al., 2012[92] Multivariate	Nonspecific language impairment (N-SLI)	United Kingdom nationwide birth cohort (N =11,383)	60 months	Vocabulary at 3 years old	NR	NR	NR	NR	Mom:↓
Mossabeb, et al., 2012[93] Multivariate	Language delay measured through number of words	Born <34 weeks in Pennsylvania hospital (N=178)	26 months	NR	NR	↑	NR	NR	NR
O'Leary, 2009[94] Multivariate	Ages and Stages Communication domain delays	Randomly Ascertained Sample of Children born to moms in Western Australia Survey of Health (RASCAL) cohort (N=1,692)	24-month survey	NR	NR	NR	0	Young maternal age during pregnancy: ↑	0
Pena et al., 2011[95] Multivariate	Risk for language impairment	Latino bilingual Pre-K in Central Texas and Northern Utah (N= 1,029)	58-68 months Older age: ↓	NR	NR	NR	NR	NR	Mom: ↓
Potijk et al., 2013[96] Multivariate	Delay communication domain score on Ages and Stages Questionnaire	Community-based sample of preterm- and term-born children, (Longitudinal Preterm Outcome Project) Netherlands (N=1,470)	4 years	NR	NR	In model but NR	Lower SES: ↑	In model but NR	NR
Pruitt, 2010[97] Univariate	Specific language impairment	African American children in Louisiana (N=161)	25-100 months	NR	↑	NR	NR	NR	Mom: 0
Reilly et al., 2007[98] Multivariate	Poorer expressive language as measured by the Communication and Symbolic Behavior Scales (CSBS) and MacArthur-Bates Communication Development Inventory (CDI)	Early Language in Victoria Study cohort, Australia (N=1,720)	24 months	NR	CSBS and CDI: ↑	CSBS and CDI: ↑	CSBS and CDI: 0	CSBS: ↑ CDI: 0	Mom, CSBS and CDI: 0

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Reilly, 2009[99] Multivariate	Stuttering onset by 3 years of age	Prospective community-ascertained cohort (the Early Language in Victoria Study) Melbourne Australia (N=1,619)	24-36 months	Higher communication and symbolic behavior scales scores at 2 years old: 0 Higher Communication Development Inventory raw vocabulary score at 2 years old: ↑	0	↑	0	NR	Mom: ↑
Reilly et al, 2013[100] Multivariate	Stuttering onset by 4 years of age	Prospective community-ascertained cohort (the Early Language in Victoria Study) Melbourne Australia (N=1,619)	4 years	Higher communication and symbolic behavior scales scores at 2 years old: ↑	0	↑	0	NR	Mom higher: ↑
Roth, 2011[101] Multivariate	Severe or moderate language based on parent report	Norwegian mother and child cohort (N = 35,135 or 36,136 depending on the analysis)	36-month followup	NR	NR	NR	NR	NR	Included in model but NR
Schjolberg, 2011[102] Multivariate	Slow language development	Norwegian mothers and child cohort (N=42,107)	18 months	NR	NR	↑	Income: ↓	0	Mom: ↓ Dad: 0
Singer et al., 2001[103] Multivariate	Speech-language development delay	Very low birthweight cohort with and without bronchopulmonary dysplasia, and controls, Cleveland, OH (N=246)	36 months	NR	NR	NR	↓	NR	NR
Tallal, Ross, and Curtiss, 1989[104] Univariate	Specific language impairment	Cases and control from San Diego, CA, longitudinal study (N=130)	48-59 months	NR	Mom: ↑ Dad: ↑ Siblings: ↑	NR	NR	NR	Mom held back and history of learning problems: ↑; Dad held back: ↑

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Tomblin et al., 1991[105] Univariate	Poor communication skills	Longitudinal cohort in Iowa concerned with early identification of children with communication problems (N=662)	30-60 months	NR	NR	↑	NR	NR	Dad: ↓ Mom: 0
Tomblin, Smith and Zhang, 1997[106] Univariate	Specific language impairment	Monolingual English-speaking kindergarteners in Iowa and Illinois (N=1,102)	Kindergarten age	NR	Mom: 0 Dad: ↑	NR	NR	NR	Mom:↓ Dad: ↓
van Batenburg-Eddes, 2013[107] Multivariate	Receptive and expressive language delay at 1.5 years, expressive language delay at 2.5 years, expressive language delay across ages	Generation R Study cohort, Netherlands (N=2,483) toddlers in the Generation R Study with neuromotor development assessment at 9-15 weeks	Assessment at mean age 1.5 years and 2.6 years	NR	NR	NR	Family income in model but NR	NR	In model but NR
Van Lierde, 2009[108] Case control: Univariate	Receptive and expressive language delay	ELBW children, matched with normal birthweight controls in Flanders, Belgium (N=24)	ELBW chronological age: 38-49 months; corrected age: 35-45 months Normal: 31-44 months	NR	NR	NR	NR	NR	NR
Weindrich et al., 2000[109] Univariate	Receptive and expressive language and articulation disorders	Mannheim Study of Risk Children cohort, Germany (N=320)	54 months	NR	NR	↑	NR	NR	NR
Whitehurst, 1991[110] Univariate	Expressive language delay	Community cohort of children living on Long Island, NY (N=117)	24-38 months	NR	0	NR	NR	NR	NR

Supplemental Table 7. Risk factors—Earlier speech and language concerns through parental education (cont.)

Author, Year	Speech and Language Outcome	Population (N)	Age	Earlier Speech and Language Concerns	Family History of Language Disorders	Male	SES	Maternal Age	Parental Education
Yliherva et al., 2001[111] Multivariate	Problems in speech production, speech perception, linguistic concepts	Birth cohort, northern Finland (N=8,276)	96 months	NR	NR	All analyses: ↑	NR	0	Mom: 0
Zambrana et al., 2014[112] Multivariate	Analysis 1: Late-onset language delays Analysis 2: Transient language delays Analysis 3: Persistent language delays	Prospective community-based sample (Children included in Norwegian Mother and Child Cohort Study) (N = 10,587)	5 years	Poorer actions and gestures composite at 1.5 years: Analysis 1: ↑ Analysis 2: ↑ Analysis 3: ↑ Poorer language comprehension composite at 1.5 years: Analysis 1: ↑ Analysis 2: ↑ Analysis 3: ↑	Late talker: Analysis 1: ↑ Analysis 2: ↑ Analysis 3: ↑ Writing and reading difficulties: Analysis 1: ↑ Analysis 2: 0 Analysis 3: ↑ Unintelligible speech: Analysis 1: 0 Analysis 2: ↑ Analysis 3: 0	Analysis 1: 0 Analysis 2: ↑ Analysis 3: ↑	Poverty Analysis 1: ↑ Analysis 2: 0 Analysis 3: 0	Mom ≤ 24 Analysis 1: 0 Analysis 2: ↑ Analysis 3: ↓ Dad ≤ 25 Analysis 1: ↓ Analysis 2: ↓ Analysis 3: 0	Mom < 5 years of college: Analysis 1: 0 Analysis 2: ↑ Analysis 3: 0 Dad < 5 years of college: Analysis 1: ↓ Analysis 2: 0 Analysis 3: ↓
Zubrick, 2007[113] Multivariate	Ages and Stages Communication domain delays	RASCAL cohort (N=1,766)	24 months	NR	↑	↑	Family income: 0 SES for neighborhood: 0	0	0

↓ = statistically significant decreased risk; ↑ = statistically significant increased risk; 0 = no statistically significant association; AC = Auditory Comprehension; CSBS= Communication and Symbolic Behavior Scales; ED = Expressive Communication; CELF-P: Clinical Evaluation of Language Fundamentals-Preschool; EC = Expressive Communication; ELBW: extremely low birth weight; G = group; N = number; N-SLI = nonspecific language impairment; NR = not reported; NY = New York; SES = socioeconomic status. SLI = specific language impairment. PLS-3: Preschool Language Scale; RASCAL = Randomly ascertained sample of children born to moms in Western Australia; SES = socioeconomic status; WILSTAAR= Ward Infant Language Screening Test Assessment Acceleration Remediation; US = United States

[a]In each study identified as reporting multivariate results, the statistical significance of each risk factor is presented controlling for all of the other identified risk factors. Unless otherwise stated, risk factors reported as NR were not included in the model.

Supplemental Table 8. Risk factors—low birthweight through other associations

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Adams-Chapman et al., 2013[77]	ELBW: 0	NR	NA: whole cohort is premature	1 month mechanical ventilation: ↑ Multiple birth: ↑	NR	NR	Cerebral palsy: ↑ Severe intraventricular hemorrhage: 0 Necrotizing enterocolitis: 0 Hearing impairment: ↑	Dysfunctional feeding:↑ Non-English speaking: ↑ Steroid exposure: 0 Black race: ↑ Private insurance: ↓
Alston, 2005[78]	NR	0	NR	NR	NR	Mother-infant time interacting: ↓ Spontaneous maternal interaction: ↓	NR	Total television: ↓ Infant babbling: ↓
Campbell, 2003[79]	NR	NR	NR	NR	NR	NR	NR	Medicaid health insurance: 0 African American race: 0
Choudhury and Benasich, 2003[80]	NR	NR	NR	NR	NR	NR	Autoimmune disease: ↑; Asthma: 0	NR
Desmarais et al., 2008[81]	NR	NR	NR	NR	↑	History of otitis media: 0	Behavior: 0 Language stimulation: 0 Lexical acquisition: 0 Communicative intent: 0 Phonetic and phonological skills: 0	NR
Everitt, Hannaford and Conti-Ramsden, 2013[82]	NR	0	NR	Mild problems:0	0	NR	Hearing concerns: 0 Ear infection: 0	Mother's occupation: 0 Father's occupation: 0
Foster-Cohen, et al., 2010[83]	NR	NR	Very preterm Receptive: ↑ Expressive: ↑	Severity of neonatal white matter abnormalities: 0	NR	Parent-child synchrony: ↓	NR	Social risk index: 0 Cognitive ability: ↓ Parent-child synchrony: ↓
Fox, Dodd and Howard; 2002[84]	NR	NR	NR	Birth difficulties:↑	NR	NR	Ear problems: 0	Sucking habits: ↑

Supplemental Table 8. Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Glascoe and Leew, 2010[85]	NR	NR	NR	NR	Elevated scores on depression screen: ↑ Anxiety: 0	Not talking to child in a special way: ↑ Not helping child learn by showing child things: ↑	NR	> 3 siblings in home ↑; ≥ 2 household moves in the past year: ↑ Limited English facility: ↑ Ethnicity: 0
Hammer, 2010[86]	NR	NR	NR	NR	NR	NR	NR	Child age: 0 Two-parent household: 0 Race/ethnicity: 0
Harrison, 2010[87]	All outcomes: 0	Older siblings: Expressive: ↑ Receptive: ↓ Vocabulary: 0	All outcomes: 0	Neonatal intensive care: 0 Breastfed >9 months: all outcomes: 0	Mom distress/ well-being: all outcomes: ↓	Support for children learning at home: Expressive and receptive: 0, vocabulary: ↓ TV watching: all outcomes: 0	Asthma: all outcomes: 0 Bronchiolitis: all outcomes: 0 Ear infections: expressive: ↑, receptive and vocabulary: 0 Ongoing hearing problems: expressive and receptive: ↑, vocabulary: 0 Social temperament: expressive and receptive: 0, vocabulary: ↓ Persistence temperament: all outcomes: ↓ Reactivity temperament: outcomes: ↑	Parents' language other than English status: expressive: ↓, receptive: 0, vocabulary: ↑ Parents' indigenous status: outcomes: 0 Number of children in household: expressive and receptive: 0, vocabulary: ↑ Smoking in household: expressive and receptive: 0, vocabulary: ↑ Neighborhood disadvantaged: all outcomes: 0
Henrichs, et al., 2011[88]	0	NR	Late bloomers: ↑		Late onset: ↑	NR	NR	Marital status: 0 Ethnicity non-western: Late bloomers: ↓ (late bloomers); Late onset: ↑ Single motherhood: Late bloomers: ↓

Supplemental Table 8. Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Kerstjens, 2011[89]	Included in model but NR	NR	Early preterm: ↑; Moderate preterm: 0	Multiple birth included in model but NR	NR	NR	NR	Non-Dutch birth included in model but NR
Kerstjens, 2009[90]	NR	NR	↑	NR	NR	NR	NR	One-parent family: ↓
Kerstjens, 2012[91]	Included in model but NR	NR	↑	Multiple birth included in model but NR	NR	NR	NR	Non-Dutch birth included in model but NR
Law, 2009[15]	SLI: 0, N-SLI: ↑	NR	NR	Mom smoked during pregnancy: SLI: 0, N-SLI: 0	NR	No reading to child: SLI: 0, N-SLI: ↑	Neurotic behaviors: SLI: 0, N-SLI: ↑ Antisocial behaviors: SLI: 0, N-SLI: ↑	No preschool: SLI ↑, Some preschool: N-SLI ↑ Mom single parent: SLI: 0, N-SLI: ↑
Law; 2012;[92]	NR	NR	NR	Small for gestational age: 0	NR	NR	NR	Pattern Construction: ↑ Behavior: ↓ Language concerns: ↓
Mossabeb, et al., 2012[93]	NR	Singleton: 0	NA	NR	NR	NR	NR	Public health insurance: ↑ Singleton gestation: 0 Small for gestational age: 0 Days on ventilator: 0 PDS ligation: 0 Culture + sepsis: 0 IVH grade 1-2: 0 IVH grade: 3-4: 0
O'Leary, 2009[94]	NR	Parity: 0	NR	Binge drinking: Prepregnancy: 0 Trimester 1:0 Trimester 2: ↑; Trimester 3: ↑	Maternal (mild): ↑	Poor parenting:↑	NR	Marital status: 0 Parent smoking: 0 Parent drug use: 0
Pena et al, 2011[95]	NR	NR	NR	NR	NR	NR	NR	Bilingual: 0 Later first English exposure: ↑

Supplemental Table 8. Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Potijk et al., 2013[96]	NR	NR	Decreasing gestational age: ↑	NR	NR	NR	NR	Multiplicative effect of SES and gestational age decreased the individual additive effect of the two associations; Number of siblings in model but NR
Pruitt 2010[97]	NR	NR	NR	NR	NR	NR	NR	NR
Reilly, 2007[98]	CSBS and CDI: 0	CSBS and CDI: 0	CSBS and CDI: 0	Twin: CSBS and CDI: 0	Mom mental health score: CSBS and CDI: 0	NR	NR	CSBS score at 12 months: ↓ Non-English-speaking background: CSBS: 0, CDI: ↑ Maternal vocabulary score: ↑
Reilly, 2009[99]	0	0	0	Twin: ↑	Mom mental health score: 0	NR	NR	Temperament: 0
Reilly et al, 2013[100]	0	Older siblings: 0	< 36 weeks: 0	Twin Birth: ↑	Mom mental health score: 0	NR	NR	Temperament: 0
Roth, 2011[101]	NR	NR	NR	Maternal use of folic acid supplements: Severe language delay: ↓ Moderate language delay: 0	NR	NR	NR	Maternal body mass index and marital status: included in models but NR

Supplemental Table 8 Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Schjolberg et al, 2011[102]	↑	NR	↑	Apgar score: ↑ Multiple birth: ↑	↑		NR	Siblings: ↑ Fussy: 0 Gestational diabetes: 0 Smoking during pregnancy: 0 Alcohol consumption first trimester: ↑ Alcohol consumption last trimester: 0 Language other than Norwegian: ↑ Daycare before 18 months: 0
Singer et al., 2001[103]	0	NR	0	Multiple birth: 0	NR	NR	Higher neurologic risk: ↑ Patent ductus arteriosis: ↑ Necrotizing enterocolitis: 0 Septicemia: 0 Peak bilirubin: 0 Retinopathy of prematurity: 0	Minority race: ↑
Tallal et al., 1989[104]	NR	NR	NR	NR	NR	NR	NR	NR
Tomblin et al., 1991[105]	0	Later: ↑	NR	NR	NR	NR	NR	At-risk determination at birth (parental background, maternal health during pregnancy, birth characteristics, and health as infant): ↑
Tomblin, Smith and Zhang, 1997[106]	0	NR	NR	C-section: 0 Duration of breast feeding: ↓	NR	NR	NR	Parent exposure to diseases, tobacco, alcohol, and drugs: 0 Maternal occupational exposure: 0

Supplemental Table 8. Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
van Batenburg-Eddes, 2013[107]	In model but NR	NR	Gestational age in model but NR	NR	NR	NR	NR	Neuromotor development: Receptive delay at 1.5 years: ↑ Receptive delay at 1.5 years: 0 Expressive delay at 2.5 years: ↑ Expressive delay across ages: ↑ Ethnicity in model but NR Marital status in model but NR
Van Lierde, 2009[108]	ELBW associated with poorer receptive language, expressive language (vocabulary, semantics, and morpho-syntaxis) and total score	NR	NR	NR	NR	NR	NR	NR
Weindrich et al., 2000[109]	NR	NR	NR	Composite measure of organic risk: ↑	NR	NR	NR	Composite measure of psychosocial risk: ↑
Whitehurst, 1991[110]	NR	NR	NR	NR	NR	NR	NR	NR
Yliherva et al., 2001[111]	Speech:↑ if low, not very low; Concepts: ↑ if very low	Production: 0 Perceptions and concepts: ↑	NR	Composite measure of risk: 0	NR	NR	Hearing impaired (all analyses): ↑	Reconstructed family: Perception and concepts: ↑ Urban residence (all analyses): 0

Supplemental Table 8. Risk factors—low birthweight through other associations (cont.)

Author, Year	Low Birthweight	Birth Order	Prematurity	Other Perinatal Factors	Parent Stress	Parenting Practices	Child Medical Conditions	Other Associations
Zambrana et al., 2014[112] Multivariate	NR	Older siblings: Analysis 1: 0 Analysis 2 (2+): ↑ Analysis 3: 0	NR	Multiple birth (all analyses): 0	NR	NR	NR	Parents with other mother tongue (all analyses): 0 Spoken to in another language (all analyses): 0 Mom partnership status (all analyses): 0
Zubrick, 2007[113]	↑	2 or more children in family: ↑	↑	Cigarette use during pregnancy: 0	Depression anxiety stress scale: 0	Parenting scale: 0 Family function: 0	NR	Paid employment: 0 Family type: 0 In daycare: 0 Other ASQ scales abnormal: gross motor: ↑ Fine motor: ↑ Adaptive score: ↑ Personal-social: ↑ Child Behavior Checklist: 0 Dimension of Temperament scale: 0

↓ = decreased; ↑ = increased;

ASQ = Ages and Stages Questionnaire; CDI = MacArthur-Bates Communication Development Inventory; ELBW = extremely low birth weight; IVH = intraventricular hemorrhage; N = number; NA = not applicable; NR = not reported in univariate analyses or included in a reported multivariate analysis, N-SLI = Nonspecific language impairment; PDS = polydioxanone, SES = socioeconomic status; SLI = specific language impairment; TV = television.