

Population-Based Screening of Children for Specific Speech and Language Impairment in Germany: A Systematic Review

Jürgen Kasper^{a, b} Julia Kreis^c Fülöp Scheibler^c Delia Möller^d Guido Skipka^c
Stefan Lange^c Olaf von dem Knesebeck^a

^aDepartment of Medical Sociology, University Medical Center Hamburg-Eppendorf, ^bUnit of Health Sciences and Education, University of Hamburg, Hamburg, ^cDepartment of Non-Drug Interventions, Institute for Quality and Efficiency in Health Care, Köln, and ^dVoice, Speech and Hearing Disorders, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Key Words

Speech and language delay • Specific language impairment • Primary language disorder • Preschool children • Screening • Systematic review

Abstract

Objective: The study was aimed at evaluating the effectiveness of a systematic population-based screening programme for specific language impairment (SLI) in preschool children in Germany. **Methods:** The study question was divided into a review of (1) evidence from studies evaluating screening programmes, (2) diagnostic instruments in the German language, and (3) studies evaluating speech and language interventions. A systematic database search was conducted between June and October 2007 and was updated in January and again in May 2008. Relevant studies were identified by 2 independent reviewers based on screened titles/abstracts and full texts. **Results:** 4,806 studies were screened. The only existing controlled screening study did not provide data for SLI. No diagnostic study met the inclusion criteria. Sixteen randomized intervention studies were included, 3 studies contributed to the appraisal of earlier

against later initiation of treatment. Most studies were of limited quality. We found indications of short-term positive effects from language therapies in children with SLI. Long-term outcomes were not investigated. No evidence supporting the advantage of earlier treatment initiation was identified. **Conclusions:** The benefit of population-based language screening of preschool children with SLI is not proven. Controlled screening studies are therefore necessary. For Germany, the accuracy of existing diagnostic instruments has not yet been sufficiently examined.

Copyright © 2011 S. Karger AG, Basel

Introduction

Acquisition of speech and language is a core process in the general development of the child. Therefore, disordered speech or language is likely to indicate a risk of a wide range of serious secondary problems during school age. This applies for instance to school performance, to reading or writing and to social and emotional quality of life. Depending on a child's age and developmental stage, the natural course of speech and language disorders im-

plicates a variety of speech- and language-related symptoms on different linguistic levels, such as an insufficient active or passive lexicon or the inability to perceive or pronounce specific phonemes. One of the earliest symptoms of the disorder is a delayed start of speech. However, late talking may also appear within the range of a normal development. Children who show significant language delay but later catch up with their peers are known in the literature as late bloomers [1, 2]. It is not clear whether those children are more at risk of language disorder than earlier developing children [3]. Diagnostically, there has to be a division between language impairment and specific language impairment (SLI) depending on whether or not the language delay appears isolated or accompanied by other developmental delay. Beyond a differential diagnosis of hearing problems, a diagnosis of SLI therefore requires the exclusion of an involvement of cognitive, motor or emotional development delays, which is difficult to diagnose at an early age.

The prevalence of language development disorder has been controversially reported. Depending on the measures and thresholds used for diagnosis, international estimations of speech and language disorders in general range from 0.65 [4] to 19% [5]. Based on the screening of more than 10,000 three-year-old children, a study in the Netherlands found a proportion of 2.4–5.3%. Data on SLI are rare. The only 2 studies explicitly reporting SLI rates found 7.4 [6] and 0.6% [7]. The latter value, however, refers to the subgroup of expressive SLI only. Likewise, the results from German studies also vary widely [8, 9]. Beyond a reliable basis for estimating the real rate of SLI, these data represent the expression of an obvious difficulty in agreeing on a diagnostic procedure, on a classification or even on a clear disease construct. A systematic review of prevalence studies by Law et al. [10] recommends avoiding the emphasis of any proxy or the averaging of existing findings. Moreover, the diversity of prevalence data indicates the need to move from the prevalence approach to a model of risk. This would involve the identification of key risk factors that are partly amenable. However, there can be no doubt about the fact that developmental speech and language disorders rank among the most frequent disorders of child development.

Few data exist on the present quality and frequency of speech and language screening in Germany. The German health system provides an opportunity to screen for developmental delay during a prevention programme conducted by paediatricians. Approximately 90% of children take part in this programme. However, language screening is not standardized in any way yet. Although there are

initiatives for language screening by the educational institutions in some of the 16 German states (*Länder*), these focus on school readiness rather than on a medical diagnosis. It can be assumed that screening procedures with proven accuracy and validity are not systematically used in any German region.

The international literature [11] suggests that children can benefit from different kinds of treatment. This holds good for children with mainly expressive symptoms more than for those with the receptive type of the disorder, for younger and older preschoolers and referring to several linguistic levels. However, evidence to justify a population-based screening programme furthermore would have to show the preponderance of a potential, additional benefit from earlier treatment initiation over potential side effects of the screening.

Our study was aimed at evaluating the evidence for the effectiveness of a systematic population-based screening programme for SLI in preschool children. This research question was processed by a group of experts on speech and language development and on systematic review methodology at the request of the German Institute for Quality and Efficiency in Health Care (IQWiG), which itself was commissioned by the Federal Joint Committee (G-BA) of the German Statutory Health Care System [12]. Evidence from existing systematic reviews such as those by Nelson et al. [13], Law et al. [10] and the National Health and Medical Research Council [14] was considered insufficient to appraise the potential benefit of screening for its implementation in the German health care system. It would have had to be proven that new studies in the field had contributed considerably to the debate in the meantime. Moreover, as existing evidence is based on international populations and diagnostic methods, it holds for English-speaking children only.

Methods

The research question predominantly implied the review of existing controlled trials, evaluating screening compared to no screening. However, if those were not identified as having sufficient explanatory power, the review had to link 2 other sources of relevant evidence [15, 16]: scientific proof of both the sufficient diagnostic quality of existing German screening procedures and the advantage of earlier compared to later initiation of treatment of SLI. Conclusions drawn from linking these 2 bodies of evidence are limited to the degree of comparability of the populations investigated in the 2 sources with each other and with the population on which the research question is focused [15].

A systematic search was conducted in bibliographic databases (AMED, BIOSIS, CCMed, CINAHL, Cochrane Central Register

of Controlled Trials CENTRAL, EMBASE including EMBASE Alert, ERIC, Heclinnet, Journals@Ovid, MEDLINE, PsycINFO, Psynindex, SciSearch, Social SciSearch) and publisher databases (Hogrefe, Karger, Kluwer, Krause & Pachernegg, Springer, Thieme). Moreover, the search strategies included the tracking of references published in existing systematic reviews, HTA (health technology assessment) reports and meta-analyses [2, 11, 13, 14, 17–22], the consultation of German experts and searches by hand in journals. The first comprehensive search was conducted between June and October 2007 and was updated in January and again in May 2008.

The studies identified by the abovementioned strategies passed through 2 rounds of assessment by 2 independent working investigators (J. Kasper, D.M.). Using titles and abstracts, the studies were screened for potential relevance regarding 1 of the 3 questions. Studies that could not be excluded with certainty were then examined in detail in full text. In cases of doubt a third investigator was consulted (F.S.). Studies meeting the inclusion criteria were assessed concerning their quality and potential biases, and data were extracted. These steps were conducted by the investigator (J. Kasper or D.M.) and reviewed twice (J. Kreis and G.S.).

To be included in the pool of studies for question 1, studies had to evaluate a screening procedure in a control group design using an unselected population. Studies contributing to question 2 had to outline accuracy data on the detection of SLI within a sample from the general population. This does not assume a language test or screening instrument could distinguish SLI from other language impairments. By use of subsequent diagnostic steps, it is nevertheless possible to determine a detection rate for SLI referring to the first test. There had to be sufficient data to generate 4-field tables, including screening positives and screening negatives and true positives as well as true negatives. For inclusion of those studies, the suitability of the reference standard was an important criterion. A reference standard was considered suitable when it was validated and not less comprehensive than the screening instrument it was compared to. Therapy studies (question 3) were included either if they evaluated a treatment method for SLI by employing a randomized controlled design with a non-treated SLI group as control, or if they allowed for comparison of effects achieved by treatment in more than 1 age group. Data were extracted from all studies meeting the inclusion criteria and summarized descriptively. All studies were rated for risk of bias, based on methods of randomization and allocation concealment, blinding, sample size calculation, comparability of groups, patient flow and intention-to-treat strategy. The study protocol was put up for public discussion. Concerns and recommendations by German experts were thoroughly scrutinized and incorporated.

Results

Screening

Out of 2,198 references, 2,031 were excluded on the basis of their title and abstract. To the remaining 167 studies another 50 studies were added, following recommendations by German experts commenting on the protocol as mentioned above. Two publications out of these

217 searched by full text were included in the review [23, 24], although they did not explicitly report results for children with SLI (fig. 1). Both referred to the same study (Rotterdam study). However, we considered this evaluation study of population screening important as an example of the kind of study that would principally allow for an appraisal of the study question.

The study was aimed at evaluating a standardized screening procedure in a particular region of the Netherlands by a controlled study. According to experience, 85% of all children participate in regular paediatric screening in Dutch child health centres. By cluster randomization, 55 of a total of 91 child health centres in the particular region were allocated to either screening by use of a standardized instrument (VTO, vroegtijdige onderkenning van ontwikkelingsstoornissen) [25] or to screening practice as usual. Paediatricians in the intervention clusters were trained in conducting the language test, whereas physicians in the control group received neither training nor test manuals. A cohort of a total of 10,942 children between 15 and 24 months old was identified that would receive regular screening twice within a specific period of a little more than 2 years. As primary endpoints, type of school, speech and language performance as well as a prognostic judgement were assessed at 8 years of age. Although it was assumed that 85% of the children would attend the regular child screening, in the study only 55% of the children randomized into the screening group had speech and language screening. As the overall response rate was unfortunately low (19–47%, depending on the outcome measure), a bias cannot be excluded. The authors report a significant effect of a 1% difference (between 2.7 and 3.7%; $p = 0.032$, one-sided testing) in school attendance between screened and unscreened children. This effect was not significant by two-sided testing. No statistically significant differences were shown between the screening and control groups based on two-sided testing at the age of 8 years. The publication does not provide an intention-to-treat analysis to prohibit potential biases due to selective dropouts. However, although there was a subgroup of 15 children with SLI within the group of children with positive screening result, no results are outlined for this group regarding the endpoints of the study [24]. Therefore, it is not clear to what extent the study results apply to the children in the focus of this review.

Diagnostics

A pool of 1,302 references, including 1,269 from database searches and another 33 which were alluded to by

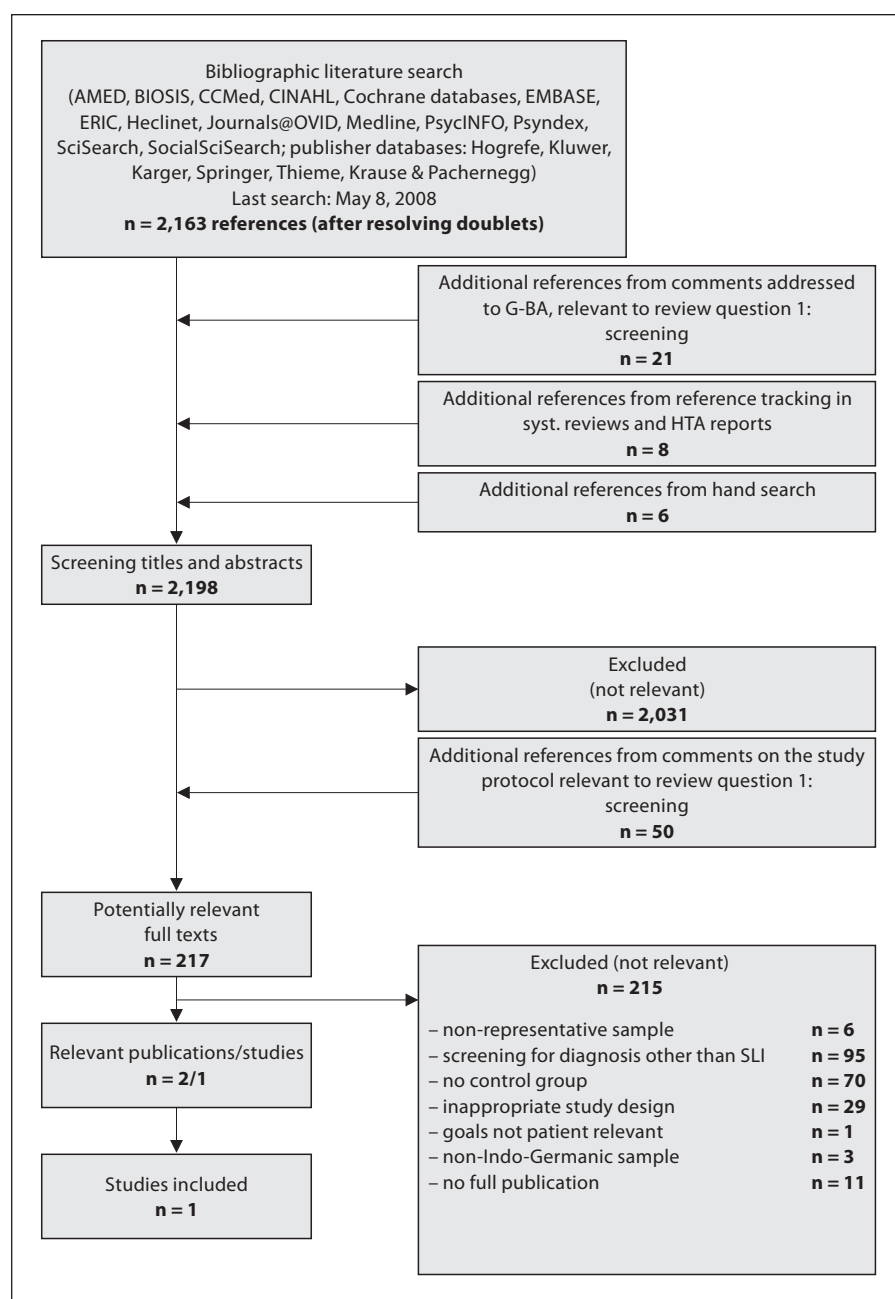


Fig. 1. Review question 1. Studies evaluating speech and language screening.

German experts or were found by reference tracking in systematic reviews, were screened based on titles and abstracts. As a result, 1,159 studies were excluded. The remaining 143 studies potentially relevant to the diagnostic review question were complemented by 79 full publications contributed by the experts (fig. 2).

None of these 222 studies screened in full text met the inclusion criteria. This means that the diagnostic quality

of the tests currently used in German-speaking countries to screen for specific speech and language development is not known. Nevertheless, since a variety of tests is in daily use by German speech and language experts, this result might need to be explained in more detail. Twenty-five publications (and 1 additional unpublished manuscript: Neumann, 2008) of diagnostic studies on 17 speech and language tests met the criterion of using any reference test

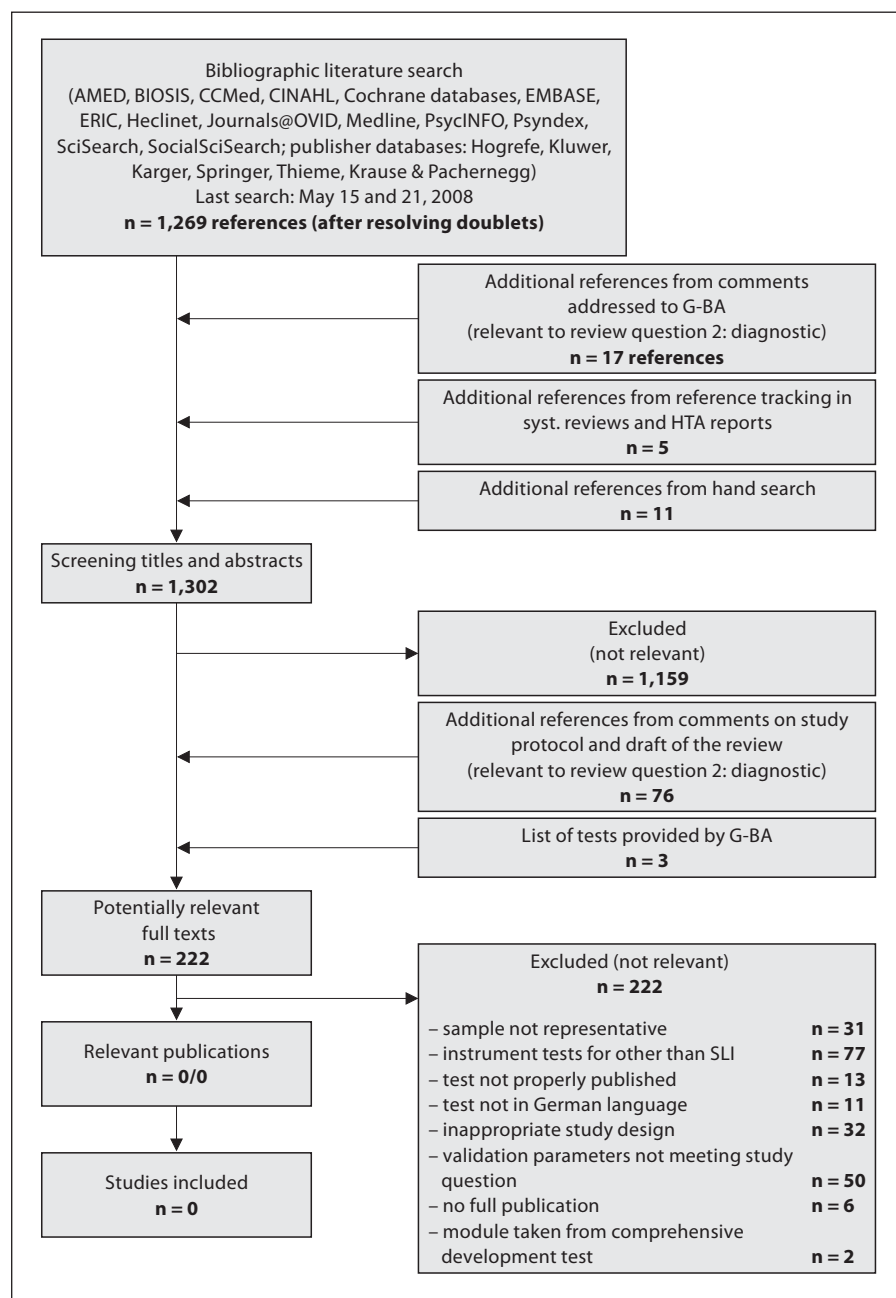


Fig. 2. Review question 2. Studies validating screening test in German language.

for proof of accuracy [1, 26–49]. Some of the reference tests used could be challenged regarding their suitability for this purpose, e.g. judgements of kindergarten teachers or results of tests that themselves are screening tests rather than comprehensive speech and language tests. Nevertheless, the difficult decision on the suitability of the reference test was superfluous since, in case of doubt, other criteria were always breached (table 1). Most of the studies

reported correlations to describe the extent of convergence with the sum score or with subscales of the reference test. This kind of data lacks information about the natural frequency of agreement between 2 tests and does not suffice to appraise test quality. Complete 4-field tables were reported for 5 tests [*Elternfragebögen für die Früherkennung von Risikokindern* (ELFRA-1, ELFRA-2); *Kindersprachscreening* (KiSS); *Sprachbeurteilung durch Eltern* –

Table 1. Methodological basic characteristics of diagnostic studies

Index test	Reference test (specific index test)	Method of validation	4-Field table provided or prepared	Accuracy of SLI detection indicated	Reference No.
AWST	kindergarten teachers' judgement	correlation	none	no	26
	complete sentences and complete words subtests from PET (Psycholinguistic Development Test) (Angermaier 1977)	correlation	none	no	
	test battery not further explained	multiple correlation	none	no	
BUEVA	WET (Vienna Development Test) (Kastner-Koller and Deimann 1998)	correlation	none	no	29
	Potsdam education questionnaire (Esser and Wyschkon 2000)	correlation	none	no	
	school years (at 8 years of age)	correlation	none	no	
ELAN	MFED 2–3 (Munich Functional Development Scales) (Köhler and Egelkraut 1994)	correlation	none	no	30
	expressive lexicon subscale of ELAN/16–26 months; AWST (Active Vocabulary Test 3–6) (Kiese-Himmel 2005)	correlation	none	no	
ELFRA-1	SETK-2 (language development test) (Grimm 2000)	correlation	none	no	31
	SETK-2 (Grimm 2000)	correlation	none	no	32
	ELFRA-2 (parent questionnaire for the early recognition of children at risk) (Grimm 2006)	sensitivity + specificity ^a	available	no	33, 49
ELFRA-2	SETK-2 and 3–5 (Grimm 2000)	correlation	none	no	31
	subscales of the BSID II (Bayley Scales of Infant Development) (Bayley 1993)	correlation	none	no	
	SETK-2 and 3–5 (Grimm 2000)	sensitivity + specificity ^a	incomplete	yes ^a	27
	subscales of the SETK 3–5 (Grimm 2001)	correlation	none	no	32
	ELFRA-3, adaptation of ELFRA-2 for 3-year-olds	sensitivity + specificity	available	no	34
	ELFRA-3	sensitivity + specificity ^a	available	no	49
	SETK-2 (Grimm 2000)	correlation	none	no	1
	SETK-2 (Grimm 2000)	sensitivity + specificity ^a	incomplete	yes	
	SETK-2 (Grimm 2000)	sensitivity + specificity ^a	incomplete	yes	
	RDLS-III (Reynell Developmental Language Scale) (Gibbard 1994)	sensitivity + specificity ^a	incomplete	yes	37
	SETK-2 (Grimm 2000)	sensitivity + specificity ^a	available	no	
	ELFRA-3	sensitivity + specificity ^a	available	no	
Questionnaires on language development	(draft of questionnaire vs.) lexicon and MLU from language sample	correlation	no	no	38
FRAKIS/ FRAKIS-K	type and frequency from language sample (FRAKIS: lexicon; questionnaire on early childhood development)	correlation	no	no	39
	MLU from language sample (FRAKIS: morphology of flexions, sentence complexity)	correlation	none	no	
	FRAKIS lexicon (FRAKIS-K)	correlation	none	no	
Screening of developmental disorders in 3-year-olds, language	development test for toddlers (Bühler and Hetzer 1966)	correlation	none	no	40
	self-developed questionnaire for parents and kindergarten teachers addressing stage of language development (no reference)	correlation	none	no	
	school marks in year 5	regression	none	no	

Table 1 (continued)

Index test	Reference test (specific index test)	Method of validation	4-Field table provided or prepared	Accuracy of SLI detection indicated	Reference No.
HASE	first to third level: school marks in German, reading, spelling, reading speed, reading comprehension	correlation	incomplete	no	27
KiSS	<i>Majority judgement within an expert panel; base rate free</i>	<i>sensitivity + specificity^b</i>	<i>yes</i>	<i>deducible^c</i>	28
	<i>Majority judgement within an expert panel; base rate fixed</i>	<i>sensitivity + specificity^b</i>	<i>yes</i>	<i>deducible^c</i>	
KISTE	LSV (Landau Language Development Test) (Götte 1976)	correlation	none	no	41
	judgement by kindergarten teachers/investigator/education questionnaire (no reference)	correlation	none	no	42
	HSET (Heidelberg Language Development Test) (Grimm and Schöler 1991)	correlation	none	no	
	LSV for preschoolers (Götte 1976)	correlation	none	no	
LSV	FTF-W (Frankfurt Test for Five-Year-Olds): lexicon (Raatz 1971)	correlation	none	no	43
	form creation subtest of the LSV; grammar subtest of the PET (Angermaier 1977)	correlation	none	no	
	ratings of communication by students (no reference)	descriptive	none	no	
MSVK	WST (lexicon test) for school beginners (Kamratowski and Kamratowski 1969)	correlation with sum score	none	no	44
	AWST (Active Vocabulary Test) (Kiese-Himmel 2005)	correlation with sum score	none	no	
	HSET (Grimm and Schöler 1991)	correlation with sum score	none	no	
	teacher judgement of language comprehension	correlation with sum score	none	no	
SBE-2-KT	ELFRA-2 (Grimm 2006)	sensitivity + specificity	available	no	45
SEV	judgement of speech therapy students	specificity ^d	none	no	46
SSV	SETK 3–5 (Grimm 2000); SSV for 3-year-old children	sensitivity + specificity	available	no	47
Teddy Test	kindergarten teacher questionnaire (no reference)	correlation	none	no	49
	Binet-Simon-Kramer Test (no reference)	correlation	none	no	
	learning test for preschoolers (no reference)	correlation	none	no	
	HAWI-VA (Hamburg-Wechsler Intelligence Test) (no reference)	correlation	none	no	

References to reference tests are in parentheses (round brackets) and are to be found in the reference studies listed in the right column (by reference No.). The table lists all 17 index tests validated by use of any reference test (form or scale of index test); text in italics refers to the unpublished status of the manuscript sent by the author of the KiSS. Expansions of study abbreviations are either provided in English equivalents or can be found in the reference studies. MLU = Mean length of utterances.

^a Values provided by the publication did not refer to representative proportions of screening positives and screening negatives. Thus, the data neither allow for a comparison with other tests nor for generalization. ^b Due to a rate of 23.5% of SLI, this reference test is debatable. ^c Due to cases of comorbidity in the screening positives, values are not reliable. ^d Reference test only for screening positives.

Kurztest (SBE-2-KT); *Sprachscreening für das Vorschulalter* (SSV)], 3 of them with 4-field tables based on representative proportions of screening positives and screening negatives. According to the detection of SLI, 4-field tables were presented for 2 of these tests (ELFRA-2, KiSS). Detection rates of tests for the wider diagnosis of language impairment cannot be assumed for SLI as a matter of course. Some others reported values for sensitivity and for specificity of the ELFRA-2 compared with the SETK-2 (*Sprachentwicklungstest für zweijährige Kinder*), ELFRA-3 and RDLS-III (Reynell Developmental Language Scale) [1, 34–37, 49]. These values resulted from applying the reference standard to all screening-positive children and to a selective sample of children who were screened negative. Since biases due to this procedure are likely and cannot be controlled, the reported values cannot be seen as representative either for the original sample or for the general population [50, 51]. An unpublished manuscript (Neumann, 2008) presented validation results for the KiSS test. Since the combination of 4 language tests (RDLS-III, *Psycholinguistische Analyse kindlicher Sprechstörungen*, *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen* and AWST, *Aktiver Wortschatztest*, for 3- to 5-year-old children) was found by the authors of this study to be unsuitable as a reference criterion, the KiSS was compared to an expert rating. However, as the authors admit, this reference was still in a premature developmental stage. For instance, the raters were given a baseline rate for positive screening before judgement. Moreover, the targeted diagnosis differed from SLI. Auditory perception and processing disorders were excluded, comorbidity was included. Therefore, data provided by the 4-field tables cannot be assumed to be valid for SLI. This means that no test could be included in the review.

Therapy

From a pool of 1,232 references including 1,172 from database searches and another 60 recommended by German experts found by reference tracking, 1,025 studies were definitely excluded on the basis of titles and abstracts. The remaining 207 potentially relevant studies were complemented by 55 publications contributed by German experts and were screened in full text. Of these, 203 had to be excluded (65 had non-representative samples, 72 pursued research questions different to ours, 53 studies were of unsuitable design, 4 used outcomes rated as not patient relevant, 2 studies did not contain any additional information, and for 7 studies no full publications were available). In all, 59 publications referring to 57 studies were identified as being relevant to answer re-

view question 3 [52–110] as they investigated the efficacy of any speech and language interventions (S+L interventions) in children with SLI; 33 of them were designed as randomized controlled trials [52–66, 69–71, 73, 75–77, 79, 81, 82, 87, 88, 90, 91, 94, 99, 101–103, 107], 24 built study groups using other strategies. Most of the studies (41) tested an intervention group against an active control group. Trials using a waiting group or sham intervention design were reported in 16 studies [52–66, 69–71]. Although less rigorous criteria were applied to the search of studies investigating the moderating influence of age on the efficacy of speech or language intervention, only 3 such studies could be identified (fig. 3) [52, 67, 68].

Trials comparing 2 treatment methods without a no-treatment study arm cannot be used for appraisal of the general efficacy of interventions. The comparison of 2 active groups provides relative efficacy information only and is therefore not suitable to make conclusions regarding efficacy in an absolute sense. Therefore, out of the pool of 57 studies, we selected the 16 using a sham intervention or waiting control group design to appraise the efficacy of S+L interventions for SLI.

The potential bias was rated as low in 2 of the 16 studies [56, 60]. Fourteen studies revealed important deviations from guidelines on the execution and reporting of randomized controlled trials [111]. Only 2 studies described appropriate randomization and concealment strategies [52, 62]. A sample size calculation was reported in 3 studies [52, 53, 61]. Out of 7 studies with a high number of dropouts, or no reported dropouts, only 1 conducted an intention-to-treat analysis [52]. In most of the studies the primary outcome was not clear.

With the exception of 2 studies [54, 63], all interventions evaluated in the 16 publications were conducted on English-speaking samples. One study reported results of S+L interventions provided to German children [54], another one to Chilean children [63]. The interventions addressed children between mean ages of 25 months [65, 53] and 13 years [56]. The study sample size ranged from 14 to 216 children. However, most of the studies included 20–30 children in total. All children had, as a minimum, problems concerning expressive language. Some studies employed more rigorous inclusion criteria such as a cutoff for the lexicon. Interventions were group therapies or single therapies at varying frequency and length over a period of 3 weeks to 12 months. With regard to the therapeutic approach and the intervention goal, the studies also differed widely. Some conducted broadly designed intervention programmes, for instance the Hanen Program for Parents [60, 61], while

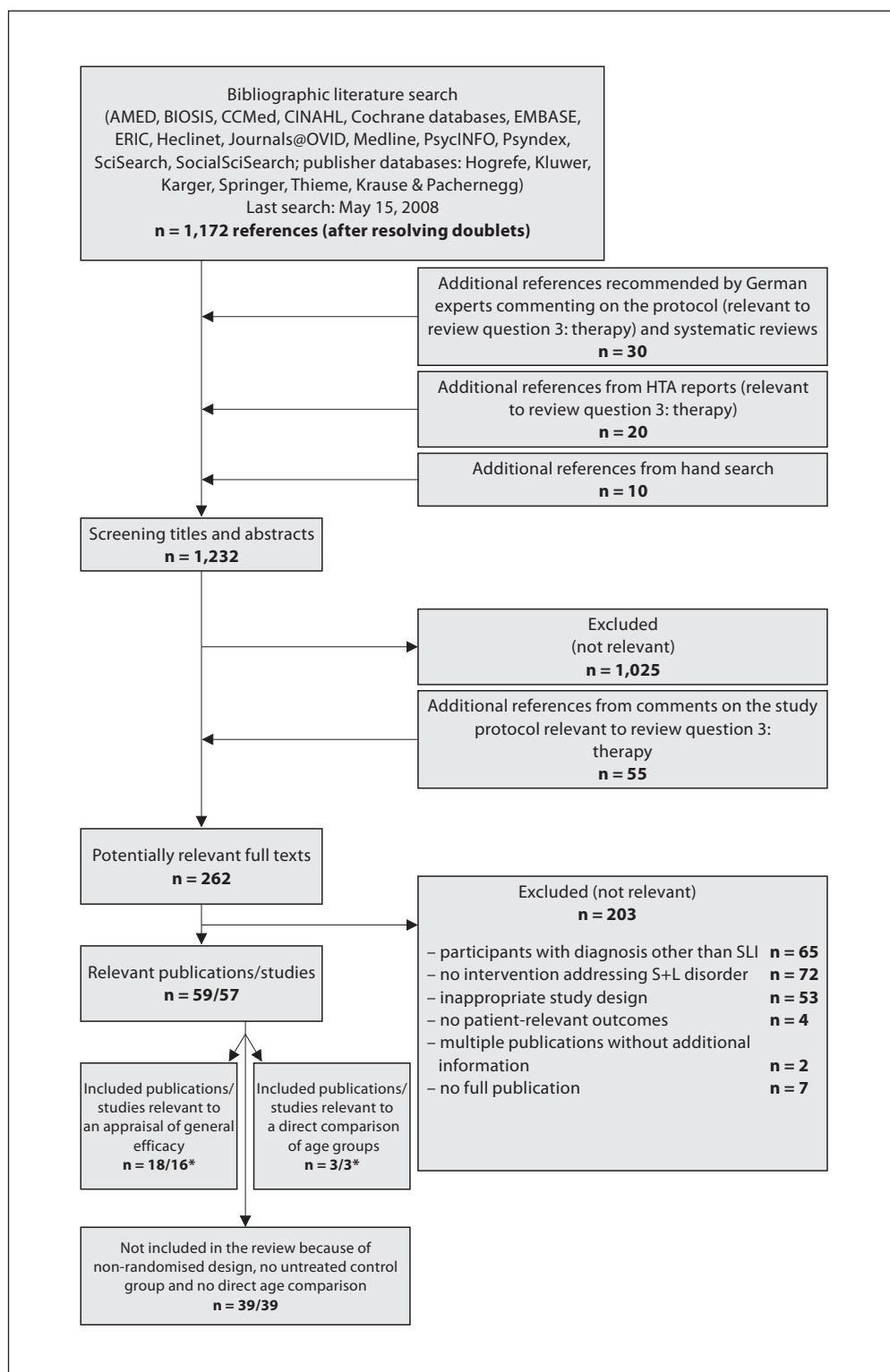


Fig. 3. Review question 3. Studies evaluating S+L interventions for children with SLI. * One study was relevant for both the question for general efficacy and the question for direct comparison of age.

Table 2. Effects of intervention studies

Linguistic level	Construct addressed	Operationalization	Reference No.	Effect	p (2-sided)
Morphology and syntax	language complexity	spontaneous language sample: mean length of utterances	52	–	0.337
			57	/	<0.001
			65	/	0.003
			71	–	>0.38
		mothers' report: mean length of utterances	57	/	<0.001
			57	/	<0.001
		spontaneous language sample: language complexity	61	–	[<0.05 ^a]
			57	/	<0.001
		Derbyshire Language Scheme Picture Test	61	/	[<0.01 ^b]
		number of linguistic markers	64	/	<0.001
	grammar	SETK-2: sentence production	53	/	0.001 ^c
		SETK 3–5: encoding of semantic relations		/	0.013 ^c
	morphology	ELFRA-2: morphology		/ /	0.017/0.021 ^{c, d}
		SETK 3–5: learning of morphological rules		–	0.376 ^c
	syntax	ELFRA-2: syntax		/ / –	<0.001/0.062 ^{c, d}
Semantics and lexicon	lexicon	language sample: different words	70	/	<0.001
			56	/	0.007 ^e
		number of different words in a game instruction	61	/	[<0.01] ^b
			65	/	<0.001
			64	/	<0.001
		language sample: total number of words	65	/	<0.001
			60	–	n.s.
			61	/	[<0.02] ^b
		ELFRA-2: expressive lexicon	65	/	<0.001
			53	/ /	0.016 ^c /0.018 ^{c, d}
				/	0.016 ^c
		SETK-2: word production		–	[0.057]
		SETK-2: comprehension I: words	57	/	<0.001
		Derbyshire Language Scheme Picture Test	57	/	<0.001
	specific aspects of lexicon	parents' report: lexicon		/	<0.001
		language sample: different target words	60	/	[≤0.02] ^b
			61	/	[<0.01] ^b
		language sample: different control words	60	–	n.s.
			61	/	[<0.01] ^b
		language sample: frequency of target words	61	/	[<0.02] ^b
			61	–	[<0.05] ^a
Phonetics and phonology	consonant production	language sample: percentage of consonants correct	52	/	0.001
			55	–	0.09
			61	–	n.g.
		language sample: inventory/position of consonants (2 operationalizations)	61	early: /	[<0.05] ^b
				medium: /	[<0.05] ^b
				late: –	[<0.05] ^a
				initial: /	[<0.05] ^b
				final: –	[<0.05] ^a
	articulation	Goldman-Fristoe Test of Articulation	52	/	0.001
		rating of phonetic inventory	69	/	<0.01
	phonological awareness	Phonological Abilities Test	55	/	<0.01
		Comprehensive Test of Phonological Processing (blending words)	58	–	n.g.
	phoneme repertoire	language sample: number of vocalizations	61	–	n.g.
	syllable structure	language sample: 3 complexity levels	61	level 1: –	n.s.
				level 2: –	n.s.
				level 3: /	[<0.01] ^b

Table 2 (continued)

Linguistic level	Construct addressed	Operationalization	Reference No.	Effect	p (2-sided)
	phoneme production	phonological test: production level, pronunciation sample, centroid	66	/	0.05
	phoneme identification	identification test	66	/	0.02
	phonological development	Assessment of Phonological Processes – Revised	52	/	0.003
		phonological error rate	62	–	0.26 ^f
Pragmatic-communicative	understandability	language sample: number of understandable utterances	57	/	0.008
		language sample: number of utterances	61	–	n.g.
		language sample: words per minute	61	–	n.g.
		total number of words	64	/	<0.001
	relevance of utterances comprehension	frequency of references to actions typical of playing	64	/	<0.001
		language sample: proportion of understandable utterances	71	–	>0.38
			65	/	<0.001
More complex constructs	narrative competence	discourse analysis	63	form: / content: /	<0.001 ^g 0.005 ^g
	general language development	RDLS	57	P: / V: /	<0.001 <0.001
		Renfrew Action Picture Test	57	I: / G: /	<0.001 <0.001
		Comprehensive Assessment of Spoken Language	58	–	n.s.
		Bristol Language Development Scales	62	–	0.73 ^d
		Preschool Language Scale	62	P: – V: /	0.44 ^f 0.025 ^f
		impairment disability	62	–	0.44 ^f
			62	–	0.56 ^f
	sentence comprehension	Token Test for Children	58	–	n.g.
		SETK-2: comprehension II: sentences	53	–	[0.323]
		SETK 3–5: understanding sentences		–	[0.113]
	expressive speech competence	SETK 3–5: both subtests ≥40		/	0.028 ^g
		SETK 3–5: min. 1 subtest ≤35		–	0.105 ^g
	literary language	Wechsler – objective reading dimension	55	R: – S: –	n.s. n.s.
		Non-Word Decoding Test	55	/	<0.05
	individual measure for improvement	compensation/non-compensation of individual study inclusion criterion	62	–	0.46 ^d
		compensation/non-compensation of any study inclusion criterion	62	/	0.036 ^d

I = Content; G = grammar; R = reading; n.g. = not mentioned; n.s. = not statistically significant; P = language production; S = writing; V = language comprehension; CDI = Communicative Development Inventory.

^a One-sided p; not significant by two-sided testing (own calculation). ^b One-sided p; also significant by two-sided testing (own calculation). ^c Declared as p from one-sided testing in the publication [53], following own calculation referring to two-sided testing.

^d Group difference at 6/12 months of follow-up. ^e Difference between all 3 groups (2 intervention groups, 1 control group) at 6/12 months of follow-up. ^f Aggregated data analysis for 6 and 12 months of follow-up. ^g Own calculation.

others focused on more specific language competences such as particular aspects of sound production or speech perception [64].

The 16 studies addressed outcomes referring to all linguistic levels: morphosyntactic, semantic and lexicon, phonetic-phonological, and pragmatic-communicative (table 2).

Effects of S+L Interventions Provided to Children with SLI on Health-Related Quality of Life

One study [62] addressed the effect of speech and language therapy on health-related quality of life using the ‘handicap’ scale of the Therapy Outcome Measures tool. No effects were shown at any measurement point between intervention and control groups.

Speech- and Language-Related Effects of S+L Interventions Provided to Children with SLI

Goal achievement concerning the field of speech and language outcomes was assessed by use of spontaneous speech samples (to evaluate the mean length of utterances, lexicon or percentage of consonants correct), by parent reports and by validated test instruments. Most of the studies employed multiple measurements, some using various operationalizations for the same parameter, such as assessing lexicon directly by a language sample and by asking the parents.

Results for the Morphosyntactic Level

Six studies addressed complexity of language, some of them by using more than one operationalization [52, 57, 61, 64, 65, 71]; 3 out of 6 showed positive effects of the intervention. Three studies addressed aspects of grammatical, syntactic and morphological ability [53, 56, 70], all of them showing positive results.

Results for the Semantic Level and Lexicon

Lexicon was assessed by 6 studies [53, 57, 60, 61, 64, 65], most of them using multiple measurements. Four studies showed clearly positive results, 2 studies showed mixed results. Girolametto et al. [59–61] investigated the use of specific words, yielding mixed results.

Results for the Phonetic-Phonological Level

Seven studies addressed parameters referring to the phonetic and phonological level, such as the percentage of correctly pronounced consonants, articulation, and either receptive or expressive phonological abilities or comprehensive phonological assessment [51, 55, 58, 61, 62, 66, 69]. Three studies showed positive, 2 studies showed mixed results. Two studies showed negative results.

Results for the Pragmatic-Communicative Level

Pragmatic measures of communication were addressed by 5 studies [57, 61, 64, 65, 71], such as the number or frequency of words within utterances [57] and the relevance or comprehensibility of utterances [71]. The results were ambiguous: 3 were positive, 2 negative.

Results for Comprehensive Language Assessments

Results for constructs referring to more than one of the abovementioned levels, such as comprehensive speech and language tests or discourse analysis [63], were reported by 6 authors. Two of them yielded positive results, 2 showed ambiguous results and 2 studies showed negative results [53, 55, 57, 58, 61, 64].

Psychosocial and Emotional Effects of S+L Interventions on Children with SLI

Three studies reported psychosocial effects of language interventions [59, 62, 64]. Using the Child Behaviour Checklist, Girolametto et al. [59] found a significant difference on the ‘externalization’ scale, but not on the ‘internalization’ scale. Control group children were described as more aggressive than children in the intervention group. However, this already applied to baseline measurement. On the Vineland Adaptive Behaviour Scales, an effect of increased social competency was found in one study [64] and not in another one [62].

Cognitive Effects of S+L Interventions on Children with SLI

Two studies [59, 62] assessed play behaviour, which was different from the control group in the study by Girolametto et al. [59] 2.5 months after the intervention, and did not differ between groups in the study by Glogowska et al. [62] 6 and 12 months after the intervention. However, no mechanism could be recognized by which the authors assumed the potential effect to be mediated. The attention level did not differ between groups in the study by Glogowska et al. [62].

Results for Earlier Compared to Later Initiation of S+L Treatment for Children with SLI

Beyond the general proof of efficacy of S+L interventions, the advantage of earlier over later initiation of such interventions is crucial for an appraisal of the benefit of screening. The related data source is quite limited [52, 67, 68]. Only one of these studies used a randomized control design [52]. In this trial an additional measurement point for both study groups was arranged after the waiting control group had received the intervention. However, the 4-month delay between the 2 groups seems too short, and a comparison of the results would have required another delay after treatment had been finished in both groups. Sommers et al. [67] investigated schoolchildren who had already passed the critical age period. Tschirner et al. [68] only provide a graphical subgroup analysis of age and therapy setting without any statistical analysis. Theoretically, indirect comparisons between age-heterogeneous study groups from different trials could be considered for appraisal as well. As a prerequisite, these pairs would have to be comparable regarding time and amount of therapy, severity of the disorder, and outcomes. This cannot be assumed for any pairs built within our pool of intervention studies. In summary, since none of the studies intended to investigate age effects, they were not designed accord-

ingly. Casual occurrence of study cohort differences regarding age turned out to be insufficient to appraise potential effects of earlier compared to later initiation of therapy.

Results for Side Effects

Side effects or undesired effects of S+L interventions such as labelling or damage of self-image were not addressed in any of the studies included in the review.

Discussion

The present study was aimed at appraising potential population-based screenings in Germany for SLI at pre-school age. Population screening would be recommendable if benefits of early detection were to outbalance any harm that inevitably arises. The position of the break-even point depends on the degree of advantage affected children can expect from earlier initiation of therapy on the one hand, and on the detection rates of true positives and true negatives on the other hand.

Our results show that neither by searching screening studies nor by linking the evidence from diagnostic and therapeutic studies a solid database can be identified to answer this question satisfactorily. The informative value of the only existing screening study for our particular research question is strongly limited by methodological weaknesses of the study [112] and by its focus on speech and language disorder in general. Furthermore, none of the available German tests reported comprehensive detection rates or could be recommended for screening. Regarding research question 3, the studies selected to appraise evidence of the efficacy of S+L interventions reported more positive than negative outcomes. However, most of the studies were only of weak quality. The justification of design-specific characteristics such as sample size, selection of particular diagnoses, and intervention strategy is not explicit in most of the studies. Since the interventions were evaluated in terms of small steps referring to particular speech and language phenomena, the relevance and sustainability of potential effects can be challenged. Long-term effects were not reported. This means that the review of our study pool hardly permits any generalizability of intervention effects. It is not clear which factors regarding characteristics of the disorder, the children or the intervention these effects refer to. The fact that the two largest and least biased studies [58, 62] showed no language-related effects implies even more cautious appraisal of the results.

Moreover, the review showed that hitherto no research has been published to answer the question on whether children with SLI would benefit from earlier initiation of a therapy. Potential side effects of screening, such as consequences of false-positive test results, were not addressed by any study.

Our search identified a couple of other systematic reviews studying similar research questions. A review of screening for speech and language disorder was intended to inform guideline development by the US Preventive Service Task Force [13, 22]. Another review focusing on the same question was conducted at the request of the National Health Service [2, 11]. According to our review, both studies concluded that, due to insufficient evidence regarding basic questions, population screening should not be recommended. This conclusion was reaffirmed by the short report of the National Screening Committee 2005 [113]. Although answering the same review question, the studies differ from ours in some detail. Due to the later appearance of the studies by de Koning et al. [23] and van Agt et al. [24] which were found by our search, neither Nelson et al. [13] nor Law et al. [10] could identify any controlled screening trials. No appraisal was made by the two reviews leading to statements about the German diagnostic instruments. However, both reviews investigated the efficacy of interventions addressing primary language disorder. Moreover, a Cochrane review focusing on this question was published by Law et al. [11] as well as a meta-analysis based on the same data [20]. Referring to the same time frame, the two other reviews did not identify an identical study pool. Due to a more rigorous application of inclusion criteria, some of the studies selected by Law et al. [10, 11, 20] and Nelson et al. [13] were excluded from our review. In particular, some studies in the Law and in the Nelson study pools were excluded from our review because of an unclear diagnosis of the children, or because a sham or waiting condition was missing. In contrast to our study, Law et al. [10, 11] included only the primary outcomes in their reviews. Notwithstanding our study, Nelson et al. [13] limited the study pool to those investigating children not older than 5 years and published in the English language. Both Law and Nelson conclude that S+L interventions are efficacious concerning language outcomes. Law et al. [10] differentiated between expressive phonological disorders and expressive syntactic and receptive types, of which only the first showed a clear benefit. With regard to the heterogeneity of the study pool, Nelson et al. [13] valued her findings of proof of efficacy as not generalizable. More-

over, they claimed missing evidence for a long-term efficacy of such therapies.

Regarding our search strategy, we consider it unlikely that relevant studies have been neglected. However, some unpublished studies were difficult to acquire. Finally, we were unable to obtain one unpublished thesis (Lancaster G: The effectiveness of parent administered input training for children with phonological disorders; MSc thesis, City University, London, 1991) identified by other systematic reviews. However, we do not assume this study would have changed our results.

As indicated by strongly diverging reports of prevalence [6, 7], the proceedings used for the diagnosis of speech and language disorders and SLI diverge. Since neither a gold standard nor a reference standard exists, study samples are selected by use of different instruments, cutoffs or expert judgements. Moreover, the distinction between primary and secondary language delay, as it is used in the international literature, implies a much less rigorous use of the diagnostic criteria compared to the definition of SLI by the ICD-10. Beyond the exclusion of evident causes of language delay such as hearing or psychiatric disorders, the ICD-10 requires the application of the discrepancy criterion, which is used to relate language delay to a child's overall development status. In a strict sense, none of the studies in the review study pool met the ICD-10 criteria for SLI. We included studies reporting a procedure for excluding hearing problems or cognitive deficits as possible reasons for language delay, although no attempts were made to apply the discrepancy criterion. This means that, as a concession to international research practice, we had to tolerate some diagnostic heterogeneity in our study pool. In a stricter sense, this limits the explanatory power of our results when applied to our target group, which was SLI.

Our results regarding the effectiveness of S+L interventions are presented on a descriptive level of analysis. They do not allow a clear estimation of the degree to which children can benefit from different kinds of S+L interventions. However, the outcomes represented in our study pool were quite heterogeneous. Moreover, outcomes appeared in varying forms of operationalization in the studies extracted, which would have made a quantitative comparison difficult. Furthermore, the studies within the pool were also diverse regarding their clinical background. A meta-analysis would therefore have been meaningless, observed differences within groups of analysis would have been difficult to attribute and genuine differences in effects would have been obscured [114, 115].

Since, on the whole, a sample size calculation was not outlined and primary endpoints were not explicit, a selec-

tive publication seems likely. Moreover, no study protocol of any of the studies in our pool was published in advance.

The study results imply a pronounced need for further research. This applies to each of the fields addressed by our study questions. Proof of efficacy of S+L interventions that can provide support for population screening has to consider the potential benefit of an earlier initiation of therapy. This requires a change to more sophisticated study designs to evaluate and compare effects of treatment of comparable disorders at earlier and later stages in the process of language development. Since a potential advantage of earlier treatment is not likely to appear and cannot be reasonably expected to be expressed in short-term outcomes as they are commonly used in speech and language evaluation trials, more sustainable and patient-relevant outcomes have to be addressed, such as, for instance, school performance. Moreover, the bigger the difference in age between two study groups, the more the law of equal treatment becomes a serious problem. On the other hand, an age-adjusted treatment limits the comparability of effects within groups. Another fundamental difficulty of studies addressing the benefit of earlier treatment consists in the comparability of study samples. Due to higher false-positive rates in samples screened at an earlier age, on average these children would be assumed to show more improvement just because of a higher percentage of spontaneous remission of speech delay. As becomes obvious, designing a reasonable intervention study for this particular purpose is not an easy task, and well-designed, controlled trials of screening programmes are necessary. However, our study shows that at present neither a screening instrument with known detection rates nor a reference standard for the validation of the latter exists in Germany. With regard to explanatory power and the extreme efforts to conduct a screening study, it is important to resolve the issue of a lack of diagnostic measurement before its initiation.

Conclusions

At present, the benefit of population-based language screening for preschool children with SLI is not proven. This does not exclude a potential benefit. However, the review shows that even prerequisites for a controlled study evaluating language screening are lacking. Research is needed in Germany to validate a diagnostic reference standard and, moreover, to validate a screening test.

Acknowledgements

This study was conducted by the Department of Medical Sociology in cooperation with the Unit of Health Sciences and Education and the Department of Voice, Speech and Hearing Disor-

ders in contract with and funded by the Institute for Quality and Efficiency in Health Care (IQWiG). We thank Sigrid Droste and Elke Hausner for the database search, Jens Klein and Makus Bönnte for ordering the literature, and Marian Cairns for editorial support.

References

- Sachse S, von Suchodoletz W: Früherkennung von Sprachentwicklungsverzögerungen mit Elternfragebogen? *Kinderarztl Prax* 2007;78:194–199.
- 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;2:1–184.
- Schlesiger C: 'Late talkers' und Prävention von Sprachentwicklungsstörungen: Früherkennung und Intervention bei spät sprechenden Kindern; in Subellok K, Bahrfeck-Wichitill K, Dupuis G (eds): *Sprachtherapie. Fröhliche Wissenschaft oder blinde Praxis*. Oberhausen, Athena, 2005, pp 206–218.
- Paul TJ, Desai P, Thorburn MJ: The prevalence of childhood disability and related medical diagnoses in Clarendon, Jamaica. *West Indian Med J* 1992;41:8–11.
- Beitchman JH, Nair R, Clegg M, Patel PG, Ferguson B, Pressman E, et al: Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *J Speech Hear Disord* 1986;51:98–110.
- Tomblin JB: A system for the diagnosis of specific language impairment in kindergarten children. *J Speech Hear Res* 1996;39:1284–1294.
- Stevenson J, Richman N: The prevalence of language delay in a population of three-year-old children and its association with general retardation. *Dev Med Child Neurol* 1976;18:431–441.
- Grimm H: Nachgefragt: Jedes vierte Kind ist sprachentwicklungsauffällig? *LOGOS Interdisziplinär* 1998;6:38–43.
- Walter M: Ergebnisse einer epidemiologischen Untersuchung zur Häufigkeit sprachlicher Förderbedürftigkeit bei Vorschulkindern in Bayern. *Sprachheilarbeit* 2007;52:146–151.
- 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;35:165–188.
- Law J, Garrett Z, Nye C: Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Syst Rev* 2003;3:CD004110.
- Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG: Früherkennungsuntersuchung auf Entwicklungsstörung des Sprechens und der Sprache bei Kindern. 2009. http://www.iqwig.de/download/S0601_Abschlussbericht_Frueherkennung_umschriebener_Stoerungen_des_Sprechens_und_der_Sprache.pdf.
- 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;117:298–319.
- National Health and Medical Research Council: Child health screening and surveillance: a critical review of the evidence. Canberra, NHMRC, 2002.
- Fryback DG, Thornbury JR: The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88–94.
- Medical Services Advisory Committee: Guidelines for the assessment of diagnostic technologies. Canberra, MSAC, 2005. [http://www.msac.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/\\$File/Diag%20Guidelines%20Sept%202005%20updated%2021%20may%202007.pdf](http://www.msac.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/$File/Diag%20Guidelines%20Sept%202005%20updated%2021%20may%202007.pdf).
- Agence Nationale d'Accréditation et d'Évaluation en Santé: L'orthophonie dans les troubles spécifiques du développement du langage oral chez l'enfant de 3 à 6 ans. Paris, ANAES, 2001.
- Cirrin FM, Gillam RB: Language intervention practices for school-age children with spoken language disorders: a systematic review. *Lang Speech Hear Serv Sch* 2008;39:S110–S137.
- Law J, Boyle J, Harris F, Harkness A, Nye C: The feasibility of universal screening for primary speech and language delay: findings from a systematic review of the literature. *Dev Med Child Neurol* 2000;42:190–200.
- Law J, Garrett Z, Nye C: The efficacy of treatment for children with developmental speech and language delay/disorder: a meta-analysis. *J Speech Lang Hear Res* 2004;47:924–943.
- Law J: The implications of different approaches to evaluating intervention: evidence from the study of language delay/disorder. *Folia Phoniatri Logop* 2004;56:199–219.
- Nelson HD, Nygren P, Walker M, Panoscha R: Screening for speech and language delay in preschool children. *Systematic Evidence Review No 41*. Rockville, Agency for Healthcare Research and Quality, 2006.
- de Koning HJ, de Ridder-Sluiter JG, van Agt HM, Reep-van den Bergh CM, van der Stege HA, Korfage IJ, et al: A cluster-randomised trial of screening for language disorders in toddlers. *J Med Screen* 2004;11:109–116.
- 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;120:1317–1325.
- Gerritsen FME: VTO language-screening 3-to 6-years olds: the development of language screening instruments to use in youth health care (in Dutch). Leiden, Rijksuniversiteit Leiden, 1988.
- Sachse S, von Suchodoletz W: Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr* 2008;29:34–41.
- Roos J, Schöler H, Treutlein A: Zur prognostischen Validität des Heidelberger Auditiven Screenings in der Einschulungsdiagnostik HASE. Abschlussbericht des Projektes EVER 2007. http://www.ph-heidelberg.de/wp/schoeler/Dateien/Abschlussbericht_EV-ER-HASE_Feb-2007.pdf (accessed May 6, 2009).
- Kiese-Himmel C: Aktiver Wortschatztest für 3- bis 5-jährige Kinder (AWST-R). Göttingen, Beltz, 2005.
- Esser G: Basisdiagnostik für umschriebene Entwicklungsstörungen im Vorschulalter (BUEVA). Göttingen, Beltz, 2002.
- Bockmann AK, Kiese-Himmel C: ELAN: Eltern Antworten. Elternfragebogen zur Wortschatzentwicklung im frühen Kindesalter. Göttingen, Beltz, 2006.
- Doil H: Die Sprachentwicklung ist der Schlüssel. Frühe Identifikation von Risikokindern im Rahmen kinderärztlicher Vorsorgeuntersuchung; thesis, Bielefeld, 2002. <http://bieson.uib.uni-bielefeld.de/volltexte/2003/250/pdf/0001.pdf>.
- Grimm H, Doil H: Elternfragebögen für die Früherkennung von Risikokindern (ELFRA-1 und ELFRA-2). Göttingen, Hogrefe, 2006.

- 33 Sachse S, Saracino M, von Suchodoletz W: Prognostische Validität des ELFRA-1 bei der Früherkennung von Sprachentwicklungsstörungen. *Klin Padiatr* 2007;219:17–22.
- 34 Hoffmann N: Die prognostische Validität einer Kurzfassung des Elternfragebogens für Zweijährige (ELFRA-2) im Vergleich zur Langfassung; MA thesis, Ludwig Maximilians University, Munich, 2007.
- 35 Sachse S, Anke B, von Suchodoletz W: Früherkennung von Sprachentwicklungsstörungen: Ein Methodenvergleich. *Z Kinder Jugendpsychiatr Psychother* 2007;35:323–331.
- 36 Sachse S, Pecha A, von Suchodoletz W: Früherkennung von Sprachentwicklungsstörungen: Ist der ELFRA-2 für einen generellen Einsatz bei der U7 zu empfehlen? *Monatsschr Kinderheilkd* 2007;155:140–145.
- 37 Walter M: Der Einsatz von ELFRA-2 und SETK-2 in einer Kinderarztpraxis: Früherfassung von Sprachentwicklungsstörungen im Rahmen der Vorsorgeuntersuchung. *Sprachheilarbeit* 2005;50:234–240.
- 38 Szagun G, Steinbrink C: Typikalität und Variabilität in der frühkindlichen Sprachentwicklung: Eine Studie mit einem Elternfragebogen. *Sprache Stimme Gehör* 2004;28:137–145.
- 39 Szagun G, Stumper B, Schramm SA: Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS) und FRAKIS-K (Kurzfassung). Frankfurt, Pearson, 2009.
- 40 Hortmann K: Zur Diagnostik des sprachlichen Entwicklungsstandes dreijähriger Kinder. *Padiatr Grenzgeb* 1987;26:149–157.
- 41 Kasielke E, Rissmann A, Scheidereiter U: Validierung des Kindersprechtests (KISTE) mit Hilfe des Landauer Sprachentwicklungstests für Vorschulkinder (LSV). *Rep Psychol* 1993;18:24–32.
- 42 Häuser D, Kasielke E, Scheidereiter U: KISTE. Kindersprachtest für das Vorschulalter. Weinheim, Beltz, 1994.
- 43 Götte R: Landauer Sprachentwicklungstest für Vorschulkinder: LSV. Ein Individualtest zur Erfassung von Artikulation, Wortschatz, Formen- und Satzbildungsfähigkeit sowie Kommunikationsfähigkeit vier- bis sechseinhalbjähriger Kinder. Weinheim, Beltz, 1976.
- 44 Elben CE, Lohaus A: Marburger Sprachverständnistest für Kinder: MSVK. Göttingen, Hogrefe, 2000.
- 45 von Suchodoletz W, Sachse S: Sprachbeurteilung durch Eltern. Kurztest für die U7 (SBE-2-KT). Handbuch. 2008. <http://www.kjp.med.uni-muenchen.de/Forschung/Sprachstufungen/SBE-2KT.php> (accessed September 8, 2008).
- 46 Heinemann M, Höpfner C: Screening-Verfahren zur Erfassung von Sprachentwicklungsverzögerungen (SEV) im Alter von 3 1/2 bis 4 Jahren bei der U8. *Kinderarzt* 1992;23:1635–1639.
- 47 Grimm H: Sprachscreening für das Vorschulalter: SSV. Göttingen, Hogrefe, 2003.
- 48 Quaiser-Pohl C: Der Teddy-Test; in Fay E (ed): *Tests unter der Lupe*. Lengerich, Pabst, 2001, pp 125–137.
- 49 Sachse S: Früherkennung von Sprachentwicklungsstörungen; in von Suchodoletz W (ed): *Früherkennung von Entwicklungsstörungen*. Göttingen, Hogrefe, 2005, pp 155–189.
- 50 Kreis J, Scheibler F, Lange S: Re: Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr* 2009;30:175–176.
- 51 Sachse S, von Suchodoletz W: Author reply. *J Dev Behav Pediatr* 2009;30:176.
- 52 Almost D, Rosenbaum P: Effectiveness of speech intervention for phonological disorders: a randomized controlled trial. *Dev Med Child Neurol* 1998;40:319–325.
- 53 Buschmann A, Jooss B, Rupp A, Feldhusen F, Pietz J, Philippi H: Parent-based language intervention for two-year-old children with specific expressive language delay: a randomized controlled trial. *Arch Dis Child* 2009;94:110–116.
- 54 Buschmann A, Jooss B, Blaschikowitz H, Koch-Graus A, Schumacher D, Heggen I, et al: Frühe Sprachförderung; in Karch D, Pietz J (eds): *Aktuelle Neuropädiatrie 2006*. Nürnberg, Novartis Pharma, 2007.
- 55 Denne M, Langdown N, Pring T, Roy P: Treating children with expressive phonological disorders: does phonological awareness therapy work in the clinic. *Int J Lang Commun Disord* 2005;40:493–504.
- 56 Ebbels SH, van der Lely HK, Dockrell JE: Intervention for verb argument structure in children with persistent SLI: a randomized control trial. *J Speech Lang Hear Res* 2007;50:1330–1349.
- 57 Gibbard D: Parental-based intervention with pre-school language-delayed children. *Eur J Disord Commun* 1994;29:131–150.
- 58 Gillam RB, Loebe DF, Hoffman LM, Bohman T, Champlin CA, Thibodeau L, et al: The efficacy of Fast ForWord language intervention in school-age children with language impairment: a randomized controlled trial. *J Speech Lang Hear Res* 2008;51:97–119.
- 59 Girolametto L, Pearce PS, Weitzman E: The effects of focused stimulation for promoting vocabulary in young children with delays: a pilot study. *J Child Commun Dev* 1996;17:39–49.
- 60 Girolametto L, Pearce PS, Weitzman E: Interactive focused stimulation for toddlers with expressive vocabulary delays. *J Speech Hear Res* 1996;39:1274–1283.
- 61 Girolametto L, Pearce PS, Weitzman E: Effects of lexical intervention on the phonology of late talkers. *J Speech Lang Hear Res* 1997;40:338–348.
- 62 Glogowska M, Roulstone S, Enderby P, Peters TJ: Randomised controlled trial of community-based speech and language therapy in preschool children. *BMJ* 2000;321:923–926.
- 63 Maggiolo M, Pavez MM, Coloma CJ: Narrative intervention for children with specific language impairment (in Spanish). *Rev Logop Foniatr Audiol* 2003;23:98–108.
- 64 Robertson SB, Weismer SE: The influence of peer models on play scripts of children with specific language impairment. *J Speech Lang Hear Res* 1997;40:49–61.
- 65 Robertson SB, Weismer SE: Effects of treatment on linguistic and social skills in toddlers with delayed language development. *J Speech Lang Hear Res* 1999;42:1234–1248.
- 66 Rvachew S: Speech perception training can facilitate sound production learning. *J Speech Hear Res* 1994;37:347–357.
- 67 Sommers RK, Schaeffe MH, Leiss RH, Gerber AJ, Bray MA, Fundrell D, et al: Effectiveness of group and individual therapy. *J Speech Hear Res* 1966;9:219–225.
- 68 Tschirner D, Hielscher-Fastabend M, Jungmann T: Relative Effektivität von Sprachfrühinterventionen bei zweijährigen Risikokindern: Zwei Programme im Vergleich. *Sprachheilarbeit* 2007;52:188–196.
- 69 Tufts LC, Holliday AR: Effectiveness of trained parents as speech therapists. *J Speech Hear Disord* 1959;24:395–401.
- 70 Wilcox MJ, Leonard LB: Experimental acquisition of wh-questions in language-disordered children. *J Speech Hear Res* 1978;21:220–239.
- 71 Yoder PJ, Camarata S, Gardner E: Treatment effects on speech intelligibility and length of utterance in children with specific language and intelligibility impairments. *J Early Interv* 2005;28:34–49.
- 72 Baxendale J, Hesketh A: Comparison of the effectiveness of the Hanen Parent Programme and traditional clinic therapy. *Int J Lang Commun Disord* 2003;38:397–415.
- 73 Boyle J, McCartney E, Forbes J, O'Hare A: A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technol Assess* 2007;11:1–158.
- 74 Camarata SM, Nelson KE, Camarata MN: Comparison of conversational-recasting and imitative procedures for training grammatical structures in children with specific language impairment. *J Speech Hear Res* 1994;37:1414–1423.
- 75 Camarata SM, Nelson KE: Treatment efficiency as a function of target selection in the remediation of child language disorders. *Clin Linguist Phon* 1992;6:167–178.
- 76 Cohen W, Hodson A, O'Hare A, Boyle J, Durrani T, McCartney E, et al: Effects of computer-based intervention through acoustically modified speech (Fast Forward) in severe mixed receptive-expressive language impairment: outcomes from a randomized controlled trial. *J Speech Lang Hear Res* 2005;48:715–729.

- 77 Courtright JA, Courtright IC: Imitative modeling as a language intervention strategy: the effects of two mediating variables. *J Speech Hear Res* 1979;22:389–402.
- 78 Crosbie S, Holm A, Dodd B: Intervention for children with severe speech disorder: a comparison of two approaches. *Int J Lang Commun Disord* 2005;40:467–491.
- 79 Dixon G, Joffe B, Bench RJ: The efficacy of visualising and verbalising: are we asking too much? *Child Lang Teach Ther* 2001;17:127–141.
- 80 Friedman P, Friedman KA: Accounting for individual differences when comparing the effectiveness of remedial language teaching methods. *Appl Psycholinguist* 1980;1:151–170.
- 81 Fudala JB, England G, Ganoung L: Utilization of parents in a speech correction program. *Except Child* 1972;38:407–412.
- 82 Gierut JA, Morrisette ML, Hughes MT, Rowland S: Phonological treatment efficacy and developmental norms. *Lang Speech Hear Serv Sch* 1996;27:215–230.
- 83 Gill CB, Klecan-Aker J, Roberts T, Fredenburgh KA: Following directions: rehearsal and visualization strategies for children with specific language impairment. *Child Lang Teach Ther* 2003;19:85–103.
- 84 Gillon GT: The efficacy of phonological awareness intervention for children with spoken language impairment. *Lang Speech Hear Serv Sch* 2000;31:126–141.
- 85 Gillon GT: Follow-up study investigating the benefits of phonological awareness intervention for children with spoken language impairment. *Int J Lang Commun Disord* 2002;37:381–400.
- 86 Glogowska M, Roulstone S, Peters TJ, Enderby P: Early speech- and language-impaired children: linguistic, literacy, and social outcomes. *Dev Med Child Neurol* 2006;48:489–494.
- 87 Hesketh A, Adams C, Nightingale C, Hall R: Phonological awareness therapy and articulatory training approaches for children with phonological disorders: a comparative outcome study. *Int J Lang Commun Disord* 2000;35:337–354.
- 88 Hesketh A, Dima E, Nelson V: Teaching phoneme awareness to pre-literate children with speech disorder: a randomized controlled trial. *Int J Lang Commun Disord* 2007;42:251–271.
- 89 Kirk C, Gillon G: Longitudinal effects of phonological awareness intervention on morphological awareness in children with speech impairment. *Lang Speech Hear Serv Sch* 2007;38:342–352.
- 90 Kot A, Law J: Intervention with preschool children with specific language impairments: a comparison of two different approaches to treatment. *Child Lang Teach Ther* 1995;11:144–162.
- 91 Kouri TA, Selle CA, Riley SA: Comparison of meaning and graphophonemic feedback strategies for guided reading instruction of children with language delays. *Am J Speech Lang Pathol* 2006;15:236–246.
- 92 Law J, Kot A, Barnett G: A comparison of two methods for providing intervention to three-year-old children with expressive/receptive language impairment. Report No. 002. London, City University, 1999.
- 93 Leonard LB, Camarata SM, Pawlowska M, Brown B, Camarata MN: The acquisition of tense and agreement morphemes by children with specific language impairment during intervention: phase 3. *J Speech Lang Hear Res* 2008;51:120–125.
- 94 Loeb DF, Armstrong N: Case studies on the efficacy of expansions and subject-verb-object models in early language intervention. *Child Lang Teach Ther* 2001;17:35–53.
- 95 McGregor KK, Leonard LB: Facilitating word-finding skills of language-impaired children. *J Speech Hear Disord* 1989;54:141–147.
- 96 Möller D, Probst P, Hess M: Durchführung und Evaluation eines Elterntrainings bei Sprachentwicklungsverzögerung. *Prax Kinderpsychol Kinderpsychiatr* 2008;57:197–215.
- 97 Mota HB, Keske-Soares M, Bagetti T, Ceron MI, Filha MG: Comparative analyses of the effectiveness of three different phonological therapy models (in Portuguese). *Pro Fono* 2007;19:67–74.
- 98 Nelson KE, Camarata SM, Welsh J, Butkovsky L, Camarata M: Effects of imitative and conversational recasting treatment on the acquisition of grammar in children with specific language impairment and younger language-normal children. *J Speech Hear Res* 1996;39:850–859.
- 99 Schwartz RG, Chapman K, Terrell BY, Prelock P, Rowan L: Facilitating word combination in language-impaired children through discourse structure. *J Speech Hear Disord* 1985;50:31–39.
- 100 Segers E, Verhoeven L: Computer-supported phonological awareness intervention for kindergarten children with specific language impairment. *Lang Speech Hear Serv Sch* 2004;35:229–239.
- 101 Sommers RK: Factors in the effectiveness of mothers trained to aid in speech correction. *J Speech Hear Disord* 1962;27:178–186.
- 102 Sommers RK, Furlong AK, Rhodes FE, Fichter GR, Bowser DC, Copetas FG, et al: Effects of maternal attitudes upon improvement in articulation when mothers are trained to assist in speech correction. *J Speech Hear Disord* 1964;29:126–132.
- 103 Teutsch A, Fox AV: Vergleich der Effektivität von artikulatorischer vs phonologischer Therapie in der Behandlung kindlicher phonologischer Störungen: Eine Pilotstudie. *Sprache Stimme Gehör* 2004;28:178–185.
- 104 Tyler AA, Lewis KE, Haskill A, Tolbert LC: Outcomes of different speech and language goal attack strategies. *J Speech Hear Res* 2003;46:1077–1094.
- 105 Ward S: An investigation into the effectiveness of an early intervention method for delayed language development in young children. *Int J Lang Commun Disord* 1999;34:243–264.
- 106 Warrick N, Rubin H, Rowe-Walsh S: Phoneme awareness in language-delayed children: comparative studies and intervention. *Ann Dyslexia* 1993;43:153–173.
- 107 Wilcox MJ, Kouri TA, Caswell SB: Early language intervention: a comparison of classroom and individual treatment. *Am J Speech Lang Pathol* 1991;1:49–61.
- 108 Wing CS: A preliminary investigation of generalization to untrained words following two treatments of children's word-finding problems. *Lang Speech Hear Serv Sch* 1990;21:151–156.
- 109 Wright SH, Gorrie B, Haynes C, Shipman A: What's in a name: comparative therapy for word-finding difficulties using semantic and phonological approaches. *Child Lang Teach Ther* 1993;9:214–229.
- 110 Zwitman DH, Sonderman JC: A syntax program designed to present base linguistic structures to language-disordered children. *J Commun Disord* 1979;12:323–335.
- 111 Moher D, Schulz KF, Altman D: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials 2001. *Lancet* 2001;357:1191–1194.
- 112 van der Ploeg CPB, Lanting CI, Verkerk PH: Should we screen for language delay in toddlers? *Pediatrics* 2008;121:1299–1300.
- 113 National Screening Committee: Child health sub-group report on speech and language delay. 2005. <http://www.library.nhs.uk/SpecialistLibrarySearch/Download.aspx?resID=88201> (accessed May 7, 2009).
- 114 Higgins JPT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions*. Oxford, Wiley, 2008.
- 115 Lenz M, Steckelberg A, Richter B, Mühlhauser I: Meta-analysis does not allow appraisal of complex interventions in diabetes and hypertension self-management: a methodological review. *Diabetologia* 2007;50:1375–1383.