

**Früherkennungsuntersuchung
auf umschriebene
Entwicklungsstörungen des
Sprechens und der Sprache bei
Kindern: Bewertung der
KiSS.2-Studie**

Arbeitspapier

Version: 1.0
Stand: 10.01.2011

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache bei Kindern: Bewertung der KiSS.2-Studie

Auftraggeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Dillenburger Str. 27

51105 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

Berichte@iqwig.de

www.iqwig.de

Schlagwörter: spezifische Entwicklungsstörungen des Sprechens und der Sprache, umschriebene Störungen des Sprechens und der Sprache, Sprachscreening, systematische Übersicht

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	iv
Abkürzungsverzeichnis	v
1 Hintergrund	1
2 Ziel der Untersuchung	2
3 Projektbearbeitung	3
4 Methoden	4
4.1 Datenquellen	4
4.2 Kriterien für den Einschluss von Studien in die Untersuchung	4
4.2.1 Population	4
4.2.2 Intervention und Vergleichsbehandlung.....	4
4.2.3 Zielgrößen.....	4
4.2.4 Studientypen	5
4.2.5 Sonstige Studiencharakteristika.....	5
4.2.6 Ein- / Ausschlusskriterien.....	5
4.3 Beschaffung zusätzlicher Informationen	6
4.4 Informationsbewertung	6
5 Ergebnisse und weiterführende Diskussion	7
6 Zusammenfassung und Diskussion	13
7 Fazit	14
8 Literatur	15
Anhang A: Dokumentation der Autorenanfragen	17

Tabellenverzeichnis

	Seite
Tabelle 1: Kriterien zur Bewertung der Ergebnisse: interne Validität und Übertragbarkeit	9
Tabelle 2: Assoziation zwischen „exklusiver klinischer Auffälligkeit“ im KiSS.2 und dem Experten-Konsensurteil in Anlehnung an [4].....	11
Tabelle 3: Dokumentation der Autorenanfragen zur KiSS.2-Studie.....	17

Abkürzungsverzeichnis

Abkürzung	Bedeutung
CONSORT	Consolidated Standards of Reporting Trials
G-BA	Gemeinsamer Bundesausschuss
ICD	International Classification of Diseases
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KiSS	Kindersprachscreening
ROC	Receiver-Operating-Characteristic
SES	Sprachentwicklungsstörung
STARD	Standards for Reporting Studies of Diagnostic Accuracy
TREND	Transparent Reporting of Evaluations with Nonrandomised Designs
UESS	umschriebene Entwicklungsstörungen des Sprechens und der Sprache

1 Hintergrund

Im Juni 2009 hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) einen Abschlussbericht zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ erstellt (Auftrag S06-01 [1]). Bei der damaligen Bewertung der diagnostischen Studien (Teilziel 2) wurde die Studie zur Evaluation des Kindersprachscreenings (KiSS) in Hessen [2] beschrieben. Die Studie konnte nicht in die Nutzenbewertung eingeschlossen werden, da sie nicht den definierten Einschlusskriterien entsprach (siehe Abschnitt 4.2). So wurde die Referenzmethode mehrfach angepasst, um unplausibel hohe Auffälligkeitsraten zu vermeiden. Deshalb wurden die Studienergebnisse zu Sensitivität und Spezifität durch die stark eingeschränkte Eignung des Referenzstandards als kaum aussagekräftig eingestuft.

Das Teilziel 2 (Bewertung diagnostischer Verfahren), dem die KiSS-Studie im Abschlussbericht S06-01 zugeordnet war, beinhaltete als ersten Schritt die Identifizierung validierter deutschsprachiger Tests zur Feststellung einer umschriebenen Entwicklungsstörung des Sprechens und der Sprache (UESS) im Altersbereich bis 6 Jahre [1]. Die KiSS-Studie zielte u. a. auf die Entdeckung von „exklusiv klinisch auffälligen“ Kindern. Dies umfasste auch Kinder, deren Sprachentwicklungsstörung mit einer Komorbidität assoziiert war. Um Aussagen über die diagnostische Güte des Instruments für Kinder mit UESS zu treffen, hätte es einer neuen Berechnung der Vierfeldertafel unter Berücksichtigung dieses Umstandes bedurft. Da die Studie allerdings keine Kinder mit isolierten auditiven Verarbeitungs- und Wahrnehmungsstörungen einschloss, hätte nur eine Aussage über eine Gruppe getroffen werden können, die den Kindern mit UESS nach ICD-10 ähnelt. Dies wäre aber im Rahmen des Abschlussberichts als akzeptabel gewertet worden.

Im Rahmen der Beratungen zur Umsetzung der Ergebnisse des Abschlussberichts wurde der Gemeinsame Bundesausschuss (G-BA) von einer Sachverständigen auf die Ergebnisse der im Studiendesign modifizierten KiSS.2-Studie hingewiesen. Der G-BA hat daraufhin das IQWiG um eine Bewertung der KiSS.2-Studie gebeten. Diese Bewertung ist Gegenstand des vorliegenden Arbeitspapiers.

2 Ziel der Untersuchung

Ziel des vorliegenden Arbeitspapiers ist die Beantwortung folgender Fragestellungen:

1. Erfüllt die KiSS.2-Studie die im Abschlussbericht S06-01 definierten Einschlusskriterien?
2. Welchen Einfluss hat die KiSS.2-Studie auf das Fazit des Abschlussberichts S06-01, vorausgesetzt, sie erfüllt die dort definierten Einschlusskriterien?

3 Projektbearbeitung

Der Verlauf des Projekts S06-01 ist im Abschlussbericht [1] beschrieben. Der G-BA hat die Unterlagen zur KiSS.2-Studie im September 2010 an das IQWiG übermittelt.

Das vorliegende Arbeitspapier wurde im IQWiG erstellt. Vorab wurde eine interne Projektskizze erarbeitet, in der das methodische Vorgehen beschrieben wurde. Nach Fertigstellung wird das Arbeitspapier an den G-BA übermittelt und 4 Wochen später im Internet veröffentlicht.

4 Methoden

4.1 Datenquellen

Für das vorliegende Arbeitspapier wurden die vom G-BA übermittelten Dokumente herangezogen. Dabei handelte es sich um den Bericht der Studienautoren an den G-BA [3] sowie die dazugehörigen Forschungsberichte [4,5]. Die Studienautorin stellte ergänzend die ebenfalls dazugehörigen Manuskripte zur Verfügung [6-8]. Zum anderen wurden die bereits für den Abschlussbericht S06-01 vorliegenden Informationen herangezogen [1]. Eine darüber hinausgehende Informationsbeschaffung in Form einer Literaturrecherche wurde für das vorliegende Arbeitspapier nicht durchgeführt.

4.2 Kriterien für den Einschluss von Studien in die Untersuchung

Nachfolgend sind die Kriterien genannt, die Voraussetzung für den Einschluss einer Studie in die Nutzenbewertung waren (Einschlusskriterien) oder zu einem Ausschluss aus der weiteren Bewertung führten (Ausschlusskriterien). Die Kriterien für den Einschluss von Studien des vorliegenden Arbeitspapiers entsprechen denen des Abschlussberichts S06-01 [1]. Da die KiSS.2-Studie ebenso wie die vorhergehend bewertete KiSS-Studie dem Teilziel 2 (Bewertung diagnostischer Verfahren) des Abschlussberichts S06-01 zuzuordnen ist, werden ausschließlich die hierfür relevanten Kriterien aufgeführt (Abschnitt 4.2.1 bis 4.2.6.).

4.2.1 Population

Eingeschlossen wurden Studien, bei denen die Durchführung des Screenings bzw. der diagnostischen Tests bei Kindern aus der Allgemeinbevölkerung bis zum vollendeten 6. Lebensjahr (Vorschulkinder) erfolgte bzw. mit entsprechend definierten Kindern aus Kontrollgruppen.

4.2.2 Intervention und Vergleichsbehandlung

Die zu prüfenden Interventionen waren diagnostische Prozeduren zur Feststellung von (u. a. umschriebenen) Entwicklungsstörungen des Sprechens und der Sprache. Die Vergleichsintervention stellte die in der jeweiligen Studie verwendete Referenzmethode dar, die auch in der Beobachtung des natürlichen Verlaufs bestehen konnte. Betrachtet wurden ausschließlich deutschsprachige Testverfahren einschließlich deutschsprachiger Versionen internationaler Testverfahren, für die eine Validierung in deutscher Sprache vorlag und die in einem einschlägigen Publikationsorgan publiziert wurden. Es wurden keine Sprachmodule aus Gesamtentwicklungstests berücksichtigt.

4.2.3 Zielgrößen

Als Zielgrößen für die Untersuchung von Teilziel 2 galten Maße zur diagnostischen Güte und Vorhersagegenauigkeit (z. B. Sensitivität / Spezifität, Likelihood Ratios, prädiktive Werte, ggf. auch Kappa-Koeffizient oder Intraklassenkorrelationskoeffizient) bzw. zur prognostischen Güte (relative Risiken, Odds Ratios oder andere geeignete Zusammenhangsmaße).

Darüber hinaus sollten Aspekte betrachtet werden, die Aufschluss über die Eignung der Testverfahren für einen Einsatz im Rahmen von Screeningprogrammen geben (z. B. Dauer der Durchführung und Auswertung, Anforderungen an die Untersucher).

4.2.4 Studientypen

Bei Studien zur diagnostischen Güte bzw. Vorhersagegenauigkeit mussten sowohl Daten für Kinder mit als auffällig als auch für Kinder mit als nicht auffällig klassifizierten Testergebnissen und entsprechenden Ergebnissen der Referenzmethode vorliegen bzw. aus den vorliegenden Daten ableitbar sein (z. B. Vierfeldertafel oder Receiver-Operating-Characteristic[ROC]-Analyse). Bei Studien zur prognostischen Güte bzw. Vorhersagegenauigkeit mussten Daten im Verlauf für Kinder sowohl mit als auffällig als auch mit als nicht auffällig klassifizierten Testergebnissen im Sinne einer (kontrollierten) Kohortenstudie vorliegen.

4.2.5 Sonstige Studiencharakteristika

Eine weitere Eingrenzung der in die Nutzenbewertung eingehenden Studien war nicht vorgesehen.

4.2.6 Ein- / Ausschlusskriterien

Voraussetzung für die Einbeziehung der KiSS.2-Studie in die Nutzenbewertung war, dass sie alle nachfolgenden Einschlusskriterien und keines der nachfolgenden Ausschlusskriterien erfüllte.

Einschlusskriterien Teilziel 2	
D1	Kinder aus der Gesamtbevölkerung bis zur Vollendung des 6. Lebensjahres (Vorschulkinder)
D2	Testverfahren zur Diagnostik der Sprachentwicklungsstörung
D3	Publikation des Testverfahrens sowie seiner Validierung in einschlägigem Publikationsorgan
D4	deutschsprachiges Testverfahren bzw. deutschsprachige Version eines anderssprachigen Testverfahrens
D5	Studientypen wie in 4.2.4 definiert
D6	Zielgrößen wie in 4.2.3 definiert

Ausschlusskriterien	
A1	Mehrfachpublikationen ohne relevante Zusatzinformation
A2	Keine Vollpublikation verfügbar ^a
A3	Sprachmodule aus Gesamtentwicklungstests

a: Als Vollpublikation gilt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT-Statements [9,10] oder, im Falle nicht randomisierter Studien, vergleichbaren Standards (z. B. TREND [11], STARD [12]) genügt und eine Bewertung der Studie ermöglichte.

4.3 Beschaffung zusätzlicher Informationen

Die Autoren der KiSS.2-Studie wurden schriftlich kontaktiert, um Fragen zur Bewertung zu klären, die anhand der vorliegenden Daten nicht beantwortet werden konnten. Die Dokumentation der Ergebnisse dieser Autorenanfragen befindet sich in Anhang A.

4.4 Informationsbewertung

Bei der Bewertung der Informationen fand ebenfalls die im Abschlussbericht S06-01 beschriebene Methodik Anwendung [1]. Im Hinblick auf die Zielsetzungen dieses Arbeitspapiers (vgl. Kapitel 2) wurde allerdings auf eine vergleichende Gegenüberstellung der Ergebnisse mit den anderen, nicht in die Nutzenbewertung einbezogenen Einzelstudien des Teilziels 2 aus dem Abschlussbericht S06-01 verzichtet.

5 Ergebnisse und weiterführende Diskussion

KiSS.2 im Kontext der vorliegenden Nutzenbewertung

Aufbauend auf der im Abschlussbericht S06-01 [1] diskutierten Studie zur Evaluation des Kindersprachscreenings in Hessen (KiSS-Studie) wurde im Jahr 2009 die KiSS.2-Studie aufgelegt. Wie sein Vorgänger ist das KiSS.2 ein Verfahren zur Sprachstandserfassung bei vier- bis viereinhalbjährigen Kindergartenkindern. Eine zentrale Aufgabe des Verfahrens ist die Entdeckung von Sprachentwicklungsstörungen (SES) – u. a. von „exklusiver klinischer Auffälligkeit“, die der Diagnose der UESS nach ICD-10 ähnelt. Wie bereits im Abschlussbericht S06-01 dargestellt, beinhaltet diese „zusätzlich die mit Komorbiditäten assoziierte Sprachentwicklungsstörung und schließt Kinder mit isolierten auditiven Verarbeitungs- und Wahrnehmungsstörungen nicht mit ein“ [1, S. 45]. Eine explizite Detektion von UESS war nicht Ziel der KiSS.2-Studie. Den Autoren zufolge kann ein Sprachscreening diese Detektion nicht leisten [3]. Auch im Abschlussbericht S06-01 wurde darauf hingewiesen, dass „von keinem einzelnen Screeningtest erwartet werden [kann], dass er sehr spezifisch Kinder mit umschriebenen Sprachentwicklungsstörungen identifiziert. Die Diagnostik der UESS bedarf eines mehrstufigen Verfahrens“ [1, S. 41].

Wenngleich KiSS.2 nicht auf die Detektion von UESS ausgelegt war, wurden nach Angaben der Autoren (vgl. Anhang A) mittels der an die Eltern und Erzieherinnen gerichteten Fragebögen sprachrelevante Komorbiditäten erfasst. Aufseiten des KiSS.2 – also des (im vorliegenden Bericht zu bewertenden) Indextests – besteht somit die Möglichkeit, die Gruppe der UESS etwas genauer einzugrenzen. Allerdings entspricht diese Art der Erfassung der Komorbiditäten nicht der von Fachgesellschaften empfohlenen klinischen Diagnosestellung, wie sie im Abschlussbericht dargestellt wurde [1, Kap. 1]. Zudem fehlten die Angaben zu Komorbiditäten für 36,3 % der Kinder.

Die Referenzmethode der KiSS.2-Studie war für die Fragestellung dieses Arbeitspapiers nicht hinreichend, da in ihr die notwendige Ausschlussdiagnostik zur Feststellung einer UESS fehlte, wie sie im Abschlussbericht S06-01 gefordert wurde [1, S. 41 + S. 144]. Demzufolge war das Einschlusskriterium D6 (Maße zur diagnostischen Güte und Vorhersagegenauigkeit; nämlich bezüglich der Entdeckung von UESS) nicht erfüllt. Die Studie konnte somit nicht in die Nutzenbewertung zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ eingeschlossen werden, zumal auch die Einschlusskriterien D2 (Testverfahren zur Diagnostik der Sprachentwicklungsstörung) und D3 (Publikation des Testverfahrens sowie seiner Validierung in einschlägigem Publikationsorgan) nicht erfüllt waren. Einen Einfluss auf das Fazit des Abschlussberichts S06-01 kann die Studie demnach nicht haben.

Weiterführende methodische Aspekte der KiSS.2-Studie

Trotz des Ausschlusses aus der Nutzenbewertung zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ wird die Studie an dieser Stelle einer weiteren – zunächst methodischen – Betrachtung unterzogen.

Der Referenztest

Die in der KiSS.2-Studie vorgestellte Referenzmethode stellt eine Weiterentwicklung des in der ursprünglichen KiSS-Studie mehrfach modifizierten Referenztests dar. Es wurde ein Experten-Konsensurteil herangezogen, das auf Basis von klar definierten Kennwerten aus drei – z. T. validierten – Testverfahren gefällt wurde. Die durch diese klaren Vorgaben gegebene Möglichkeit einer Replikation des Referenztests deutete auf eine gewisse Standardisierung hin. Allerdings ist der Referenztest in der aktuellen Form zum ersten Mal eingesetzt worden, weshalb eine allgemeine Akzeptanz bislang nicht angenommen und er somit nicht als Referenzstandard bezeichnet werden kann. Für einen derartigen Test wäre zu fordern, dass er seinerseits einer gründlichen Validierung unterzogen wurde, im Rahmen derer adäquate Werte für Sensitivität und Spezifität hätten nachgewiesen werden müssen.

Insbesondere die Erfassung der Validität stellt in Ermangelung eines etablierten Referenz- bzw. Goldstandards eine große Herausforderung dar. Vorstellbar wäre in diesem Zusammenhang eine ausreichend lange Nachbeobachtungszeit aller teilnehmenden Kinder. Spätestens in einem Alter von 6 bis 8 Jahren sollte sich herauskristallisiert haben, welche Kinder tatsächlich von einer UEES betroffen sind und mit hoher Wahrscheinlichkeit langfristig Entwicklungsprobleme haben werden. Auch diese Methode ist nicht frei von Verzerrungsquellen. Es ist zum Beispiel vorstellbar, dass ein sogenanntes Behandlungsparadox (treatment paradox) auftritt, wenn auf Basis der Ergebnisse des zu evaluierenden Tests eine effektive Behandlung durchgeführt wird, der Referenztest aber erst später – nachdem die Behandlung begonnen oder durchgeführt wurde – erfolgt [13]. Dennoch scheint diese Methode derzeit die verlässlichste zu sein, um das Vorliegen einer UEES zu verifizieren.

Andere methodische Aspekte

Während nicht auszuschließen ist, dass der Referenztest zu systematischen Verzerrungen innerhalb der Untersuchung führte, sind andere Kriterien, die Einfluss auf die interne Validität nehmen, Autorenangaben zufolge adäquat berücksichtigt worden (vgl. Tabelle 1). Auf Nachfrage haben die Autoren die Studiendurchführung näher beschrieben (vgl. Anhang A).

Die KiSS.2-Studie war als diagnostische Querschnittstudie angelegt. In solchen Studien führt eine möglichst kurze Zeitspanne zwischen Index- und Referenztest zu einer Minimierung systematischer Verzerrungen (z. B. oben erwähntes Behandlungsparadox). Da beide Tests Autorenangaben zufolge direkt nacheinander durchgeführt wurden, wurde die Zeitspanne für das gewählte Studiendesign somit als adäquat bewertet. Das oben beschriebene Problem des Referenztests – kein Goldstandard, keine Validierung – impliziert jedoch, dass eigentlich ein anderes nämlich longitudinales Studiendesign mit einer längeren Nachbeobachtungszeit – und somit einer längeren Zeitspanne zwischen beiden Tests – hätte gewählt werden müssen (prognostic accuracy study). In einem solchen Design wäre die Zeitspanne zwischen den dann zwei Indextests (z. B. KiSS.2 und Expertenbeurteilung) weiterhin minimal, während der Referenztest notwendigerweise zeitversetzt erhoben würde.

Tabelle 1: Kriterien zur Bewertung der Ergebnisse: interne Validität und Übertragbarkeit

	Ja	Nein	Unklar
Interne Validität			
Referenztest akzeptabel?		x	
Adäquate Zeitspanne zwischen Index- und Referenztest?	x		
Interdependenz der Tests adäquat berücksichtigt?			x
Partielle Verifikation vermieden?	x		
Differenzielle Verifikation vermieden?	x		
Inkorporationsbias vermieden?	x		
Verblindung bei Interpretation der Testergebnisse?	x		
Intention-To-Diagnose-Analyse adäquat?	x		
Ergebnisunabhängige Berichterstattung?	x		
Keine sonstigen Aspekte, die das Verzerrungspotenzial beeinflussen?		x ^a	
Übertragbarkeit			
Patientenspektrum repräsentativ?		x	
Index- und Referenztest übertragbar?			x
Entscheidungsprozess bei der Testinterpretation übertragbar?	x		
Hintergrundinformationen übertragbar?	x		
Vergleichbarer Indextest in Studie und Versorgungskontext?	x		
Untersucherqualifikation vergleichbar?			x
^a Index- und Referenztest wurden von derselben Person erhoben			

Die Frage nach einer adäquaten Berücksichtigung möglicher Interdependenzen der Tests (z. B. Ermüdungserscheinungen oder Lerneffekte) war nicht zu beantworten. Die Testreihenfolge wurde zwar randomisiert. Dennoch war nicht auszuschließen, dass eine gegenseitige Beeinflussung der Tests vorhanden war, die zu einer Verzerrung der Ergebnisse führte.

Ein weiteres Kriterium, das adäquat berücksichtigt wurde, ist die partielle Verifikation. Sie wurde vermieden, indem sowohl Index- als auch Referenztest bei allen Kindern durchgeführt wurden. Auch eine differenzielle Verifikation war nicht anzunehmen, da unabhängig vom Resultat des zuerst durchgeführten Tests bei allen Kindern der jeweils andere Test durchgeführt wurde. Ein Inkorporationsbias konnte ausgeschlossen werden, da der Indextest nicht Teil des Referenztests war. Des Weiteren wurde von den Autoren schriftlich mitgeteilt, dass die Interpretation von Referenz- und Indextest ohne Kenntnis des Ergebnisses des jeweils anderen Tests erfolgte.

Die Intention-To-Diagnose-Analyse wurde als adäquat bewertet, da die im Forschungsbericht auftauchenden fehlenden Werte gering waren und von den Autoren auf Nachfrage plausibel erklärt wurden (vgl. Anhang A). Sie sind offensichtlich zufällig aufgetreten, weswegen anzunehmen war, dass sie keinen Einfluss auf die Validität der Ergebnisse hatten. Der hohe Anteil fehlender Werte zu Komorbiditäten (36,3 %; vgl. Seite 7) floss in die Bewertung des Intention-To-Diagnose-Prinzips nicht mit ein. Sämtliche Informationen zu Komorbiditäten (z. B. Erhebungsmethode, Art der Komorbidität) wurden dem Institut von den Autoren auf Nachfrage zur Verfügung gestellt. Die Erhebung der Komorbiditäten diente nicht Studienzwecken, so dass keine Auswertungsabsicht seitens der Autoren angenommen werden konnte. Die Erfassung der Komorbiditäten verfolgte den Zweck, den Sprachexpertinnen in den Gesundheitsämtern Hinweise zu liefern, ob weiterführende Untersuchungen bzw. Interventionen notwendig waren (vgl. Anhang A). Trotz dieser positiven Bewertung der Intention-To-Diagnose-Analyse innerhalb des auf SES ausgerichteten Forschungsberichts stellte der hohe Anteil fehlender Werte zu Komorbiditäten für das Ziel dieses Arbeitspapiers – das Erkennen von UESS – einen groben Mangel dar.

Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung wurden keine bemerkt. Ein letzter Aspekt, der das Verzerrungspotenzial hätte beeinflussen können, war die Erhebung von Index- und Referenztest. Die Testreihenfolge wurde zwar randomisiert, und es wurden immer derselbe Index- und Referenztest durchgeführt. Allerdings wurden beide von demselben Studenten erhoben. Es bleibt somit unklar, ob der jeweils zweite Test in seiner Durchführung unabhängig vom Resultat des zuerst durchgeführten Tests war. Insbesondere im Zusammenhang mit der gewählten Referenzmethode war das Verzerrungspotenzial der Ergebnisse der KiSS.2-Studie als hoch einzustufen.

Übertragbarkeit der Ergebnisse

Auch die Übertragbarkeit der Ergebnisse der KiSS.2-Studie auf die Gesamtbevölkerung Deutschlands ist fraglich (vgl. Tabelle 1). Zum einen war das untersuchte Kollektiv der KiSS.2-Studie nicht repräsentativ. Neben einem unausgewogenen Geschlechterverhältnis wurden ausschließlich Kinder aus Frankfurt und Offenbach untersucht. Durch die Auswahl dieser Städte lag der Anteil an Kindern mit Migrationshintergrund bei 58,5 %¹ – dem Dreifachen des bundesdeutschen Durchschnitts [14]. Zum anderen haben innerhalb der KiSS.2-Studie Studenten der Linguistik und der Logopädie sowohl den Referenz- als auch den Indextest durchgeführt. Letzterer soll in der Anwendungssituation von Erzieherinnen ausgeführt werden. Ob der in der Studie von Studenten durchgeführte Indextest dem Indextest in zukünftigen Anwendungssituationen entspricht und ob die Untersucherqualifikation innerhalb der Studie mit derjenigen im Versorgungskontext vergleichbar ist, musste aus diesem Grund infrage gestellt werden. Eine weitere Frage zur Übertragbarkeit fokussierte auf

¹ In der Autorenantwort vom 02.11.2010 (vgl. Anhang A) wurde darauf aufmerksam gemacht, dass sich im Forschungsbericht [4] auf Seite 5 ein Fehler in der Angabe zu den Probanden befindet. Demnach teilen sich die Kinder in 129 muttersprachlich deutsche und 182 muttersprachlich anderssprachige Kinder mit Migrationshintergrund auf (anstelle von 140 und 161 Kindern). Der hier angegebene Prozentwert bezieht sich auf die korrigierten Angaben.

die Vergleichbarkeit des in der Studie eingesetzten mit dem im späteren Versorgungskontext einzusetzenden Indextests – d. h. wurde die Testbatterie des Indextests seit Studiedurchführung verändert. Im Forschungsbericht wurde erwähnt, dass das dort beschriebene KiSS.2 eine Studienversion war, die in der Nachfolgezeit noch einmal leicht modifiziert wurde. Auf Nachfrage wurde von den Autoren mitgeteilt, dass die Veränderungen nicht den Test selbst – also den Screeningbogen – betrafen, sondern lediglich die an Eltern und Erzieherinnen gerichteten Fragebögen. Demnach enthielt die Studienversion zusätzliche Fragen für Studienzwecke, die im Routineeinsatz nicht von Relevanz sind (vgl. Anhang A). Die Veränderungen konnten somit als unwesentlich betrachtet werden. Eine Übertragbarkeit der Entscheidungsprozesse bei der Interpretation der Testergebnisse sowie der in der Studie vorliegenden Hintergrundinformationen konnte angenommen werden.

Eine auf den dargestellten Fakten basierende generelle Einschätzung der Übertragbarkeit der Studienergebnisse kommt zu dem Ergebnis, dass die Wahrscheinlichkeit groß ist, dass im Versorgungskontext der Bundesrepublik Deutschland keine vergleichbaren Ergebnisse erzielt würden.

Bestimmung von UESS

Trotz dieser generellen methodischen Einschränkungen der KiSS.2-Studie und deren Ausschluss aus der Nutzenbewertung zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ werden die Studienergebnisse im Hinblick auf eine Ermittlung von UESS an dieser Stelle einer näheren Betrachtung unterzogen.

Zur Quantifizierung der Übereinstimmung von Referenz- und Indextest wurden in der aktuellen Untersuchung ebenso wie in der Vorgängerstudie Vierfeldertafeln mit Angaben zur Sensitivität und Spezifität des KiSS.2-Tests geliefert – u. a. in Bezug auf die Entdeckung „exklusiv klinischer Auffälligkeit“ als Teilkategorie einer SES (siehe Tabelle 2).

Tabelle 2: Assoziation zwischen „exklusiver klinischer Auffälligkeit“ im KiSS.2 und dem Experten-Konsensurteil in Anlehnung an [4]

„Exklusive klinische Auffälligkeit“		Experten-Konsensurteil		
		Auffällig	Unauffällig	Gesamt
KiSS.2	Auffällig	20 (6,6 %)	29 (9,5 %)	49 (16,1 %)
	Unauffällig	10 (3,3 %)	245 (80,6 %)	255 (83,9 %)
	Gesamt	30 (9,9 %)	274 (90,1 %)	304 (100 %)
Sensitivität: 66,7 %; Spezifität: 89,4 %				

Unter Kenntnis und Ausschluss derjenigen Kinder, die eine sprachrelevante Komorbidität aufweisen, hätte die in Tabelle 2 dargestellte Vierfeldertafel zur „exklusiven klinischen Auffälligkeit“ neu berechnet werden können, um die diagnostische Güte für UESS zu ermitteln. Den Autoren zufolge wurden sprachrelevante Komorbiditäten mit dem Index- also dem KiSS.2-Test erfasst (vgl. Anhang A). Es sprachen jedoch mehrere Gründe gegen eine Verwendung dieser Daten und damit gegen eine neue Berechnung der dargestellten Vierfeldertafel zur Ermittlung der diagnostischen Güte für UESS:

- Wie bereits oben dargestellt, fehlten im KiSS.2-Test bei mehr als einem Drittel der Kinder Angaben zu möglichen sprachrelevanten Komorbiditäten.
- Darüber hinaus entsprach die Erfassung der Komorbiditäten mittels Eltern- und Erzieherinnenfragebögen im KiSS.2-Test nicht der von Fachgesellschaften geforderten klinischen Diagnosestellung².
- Zuletzt fehlte aufseiten des Referenztests gänzlich die Ausschlussdiagnose bezüglich sprachrelevanter Komorbiditäten, weswegen sich das Experten-Konsensurteil weiterhin auf „exklusive klinische Auffälligkeit“ bezieht und somit weder 100 % sensitiv noch spezifisch für UESS sein kann.

² In der ursprünglichen KiSS-Studie wurden die im Indextest ermittelten „exklusiv klinisch auffälligen“ Kinder einer anschließenden Konfirmationsdiagnostik unterzogen. Diese führte zu dem Resultat, dass unter den „exklusiv klinisch auffälligen“ Kindern 2 mit einer sprachrelevanten Komorbidität waren [2]. Im Gegensatz zu dieser Diagnosestellung in der KiSS-Studie führte die Erfassung von Komorbiditäten mittels Fragebögen in der KiSS.2-Studie dazu, dass sich die 6 Kinder mit sprachrelevanten Komorbiditäten auf alle Zellen der Vierfeldertafel verteilten. Per definitionem setzt eine „Ko“-Morbidität jedoch eine Grund- beziehungsweise Primärerkrankung – hier die Sprachentwicklungsstörung – voraus. Deshalb hätten Kinder mit Komorbiditäten – so wie in der ursprünglichen KiSS-Studie – eigentlich nur in den KiSS.2- Gruppen auftreten dürfen, in denen der Verdacht auf eine SES und damit eine Primärerkrankung vorlag.

6 Zusammenfassung und Diskussion

Die KiSS.2-Studie konnte ebenso wie ihre Vorgängerstudie nicht in die Nutzenbewertung zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ eingeschlossen werden. Einen Einfluss auf das Fazit des Abschlussberichts S06-01 hat sie folglich nicht. Wie die KiSS-Studie erfüllte auch die KiSS.2-Studie nicht das Einschlusskriterium D6 (Maße zur diagnostischen Güte und Vorhersagegenauigkeit; nämlich bezüglich der Entdeckung von UESS). Sie war auf die Identifikation von SES und nicht auf die Entdeckung von UESS angelegt. Aus diesem Grund fehlte innerhalb der Studie die für die Identifikation von UESS notwendige Ausschlussdiagnostik für Komorbiditäten im Referenztest. Im Indextest wurden Komorbiditäten erhoben, allerdings lediglich mittels Befragung der Eltern und Erzieherinnen. Der für die Entdeckung von UESS notwendige Ausschluss von Komorbiditäten kann den Autoren zufolge „nicht Ergebnis eines Screenings sein, sondern bedarf einer dezidierten fachärztlichen Diagnostik“ [3, S. 6]. Aus diesem Grund schreiben sie in ihrer Antwort an das IQWiG, „es wäre ratsam, wenn das IQWiG dem Rechnung tragen würde und das KiSS.2 nicht deshalb als Screening ausschließt, da es auf SES screent und nicht auf UESS“ (vgl. Anhang A). Diesem Rat konnte innerhalb dieses Arbeitspapiers nicht gefolgt werden, da der G-BA um eine Prüfung der KiSS.2-Studie im Rahmen der Beratungen zur Umsetzung der Ergebnisse des Abschlussberichts S06-01 gebeten hatte. Im Abschlussbericht wurde bereits adressiert, dass der Fragestellung diese Schwierigkeit inhärent war [1, S. 153].

Darüber hinaus bleibt zur KiSS.2-Studie festzuhalten, dass das weiterentwickelte Referenzverfahren seine Validität und Akzeptanz innerhalb der wissenschaftlichen Gemeinschaft erst noch unter Beweis stellen muss. Es wäre wünschenswert, wenn alle Kinder der KiSS.2-Studie bis zum Alter von 6 bis 8 Jahren nachbeobachtet würden. Auf diese Weise könnte im deutschsprachigen Raum erstmals ein detailliert entwickelter Referenztest zu Sprachentwicklungsstörungen validiert werden. Auch die diagnostische Güte des KiSS.2 könnte dadurch evaluiert werden. Würden Referenz- und Indextest ferner um eine – den Empfehlungen der Fachgesellschaften entsprechende – Ausschlussdiagnostik ergänzt werden, wäre zukünftig ein Einsatz zur Ermittlung von UESS denkbar. Einzig die Frage nach der Übertragbarkeit der Ergebnisse auf den Versorgungskontext der Bundesrepublik Deutschland bliebe noch offen.

7 Fazit

Da in der KiSS.2-Studie die Ausschlussdiagnostik zur Feststellung einer UESS nicht ausreichend war, ließ sich die diagnostischen Güte und Vorhersagegenauigkeit des dort eingesetzten (Index-)Tests bezüglich der Entdeckung von umschriebenen Entwicklungsstörungen des Sprechens und der Sprache nicht sicher ermitteln. Die KiSS.2-Studie erfüllte daneben weitere Einschlusskriterien für die Nutzenbewertung zum Thema „Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ (IQWiG-Abschlussbericht S06-01) nicht. Sie konnte somit nicht in die Nutzenbewertung eingeschlossen werden und hat demnach auch keinen Einfluss auf das Fazit des Abschlussberichts S06-01.

8 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Früherkennungsuntersuchung auf umschriebene Entwicklungsstörungen des Sprechens und der Sprache: Abschlussbericht; Auftrag S06-01 [online]. 17.06.2009 [Zugriff: 18.08.2009]. (IQWiG-Berichte; Band 57). URL: http://www.iqwig.de/download/S06-01_Abschlussbericht_Fruherkennung_umschriebener_Stoerungen_des_Sprechens_und_der_Sprache.pdf.
2. Neumann K, Euler HA. Bericht an das IQWiG: Katzensgoldstandards in der Sprachstandserfassung: Sensitivität und Spezifität des Kindersprachscreenings (KiSS) [unveröffentlicht]. 2008.
3. Neumann K, Euler HA. Bericht an den Gemeinsamen Bundesausschuss zur Studienlage zum Sprachstandserfassungsverfahren für vier- bis viereinhalbjährige Kinder Kindersprachscreening, Version 2 (KiSS.2) [unveröffentlicht]. 2010.
4. Neumann K, Euler HA. Einführung einer flächendeckenden Sprachstandserfassung in Hessen: Forschungsbericht an das Hessische Ministerium für Arbeit, Familie und Gesundheit für das Jahr 2009 [unveröffentlicht]. 2010.
5. Neumann K, Euler HA. Einführung einer flächendeckenden Sprachstandserfassung in Hessen: Forschungsbericht an das Hessische Ministerium für Arbeit, Familie und Gesundheit für das Jahr 2008 [unveröffentlicht]. 2009.
6. Euler HA, Holler-Zittlau I, von Minnen S, Sick U, Dux W, Zaretsky Y et al. Psychometrische Gütekriterien eines Kurztests zur Erfassung des Sprachstands vierjähriger Kinder. HNO 2010; 58(11): 1116-1123.
7. Neumann K, Euler HA. Wirksamkeitsbewertung eines Beobachtungsverfahrens zur Sprachstandserfassung für Vier- bis Viereinhalbjährige in Kindergärten [unveröffentlicht]. 2010.
8. Neumann K, Holler-Zittlau I, Van Minnen S, Sick U, Zaretsky Y, Euler HA. Katzensgoldstandards in der Sprachstandserfassung: Sensitivität und Spezifität des Kindersprachscreenings (KiSS) [eingereichtes Manuskript]. 2010.
9. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008; 148(4): 295-309.
10. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869.

11. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004; 94(3): 361-366.

12. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CT, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003; 49(1): 1-6.

13. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; 140(3): 189-202.

14. Statistisches Bundesamt Deutschland. Migration und Integration [online]. [Zugriff: 25.11.2010]. URL:

<http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Statistiken/Bevoelkerung/MigrationIntegration/MigrationIntegration.psml>.

Anhang A: Dokumentation der Autorenanfragen

Tabelle 3: Dokumentation der Autorenanfragen zur KiSS.2-Studie

Adressat / Datum	Inhalt der Anfrage	Antwort (Antwort durch / Datum / Inhalt)
Neumann, K. / 15.09.2010	Bitte der Rechtseinräumung für die Nutzung vertraulicher Dokumente	Neumann, K. / 15.09.2010: Zusendung der Rechtseinräumung
Neumann, K. / 15.10.2010	Ausführliche Anfrage nach genaueren Informationen zur Nachbeurteilung von Kindern durch eine Sprachexpertin mittels itemspezifischem Algorithmus	Neumann, K. / 02.11.2010 bzw. 05.11.2010: Ausführliche Beantwortung der Fragen inkl. detaillierter Beschreibung des itemspezifischen Algorithmus, dessen Einsatz und der resultierenden Ergebnisse
	Detaillierte Anfrage zu sprachrelevanten Komorbiditäten: Wurden diese erfasst? Wenn ja, wie und durch wen? In welchen Feldern der Vierfeldertafel zu SES befinden sich ggf. Kinder mit Komorbiditäten?	Detaillierte Beantwortung der Anfrage: Die Komorbiditäten wurden in Eltern- und Erzieherinnenfragebögen nachgefragt; es erfolgte eine Beschreibung der erfassten Komorbiditäten sowie generelle Anmerkungen zu Sprachscreenings; darüber hinaus wurde die Vierfeldertafel zu SES mit Kennzeichnung der jeweils in den Zellen befindlichen Kinder mit Komorbiditäten geliefert
	Frage nach näheren Angaben zur Tabelle 5 der aktuellen Studie [4]	Die absoluten Häufigkeiten und Stichprobenumfänge wurden ergänzt – dabei wurde darauf aufmerksam gemacht, dass die Personenzahlen der Probanden in der aktuellen Studie [4] fehlerhaft waren und es heißen muss: 129 muttersprachlich deutsche, 182 muttersprachlich anderssprachige Kinder mit Migrationshintergrund; zudem wurde die Variable „Mitarbeit des Kindes“ näher beschrieben
	Fragen zu Abweichungen im Stichprobenumfang zwischen verschiedenen Tabellen sowie zur Teilnehmerrate	Die Abweichungen im Stichprobenumfang wurden durch gelegentlich fehlende Daten begründet, die unsystematisch zustande kamen; die Teilnehmerrate lag bei nahezu exakt 80 % – die Eltern aller vier- bis viereinhalbjährigen Kinder einer Kindertagesstätte wurden kontaktiert, nur die Kinder mit unterschriebener Einverständniserklärung wurden in die Studie aufgenommen

(Fortsetzung)

Tabelle 3: Dokumentation der Autorenanfragen zur KiSS.2-Studie (Fortsetzung)

Adressat / Datum	Inhalt der Anfrage	Antwort (Antwort durch / Datum / Inhalt)
	Fragen zu den Sprachexperten bzgl. möglicher Überschneidungen mit Sprachexperten aus der KiSS-Studie, bzgl. ihrer Qualifizierung und bzgl. einer eventuellen Vorgabe zum Prozentsatz der als auffällig zu beurteilenden Kinder (z. B. 10 %)	Eine Sprachexpertin war auch Bewerterin in der KiSS-XL-Studie, der Vorläuferstudie zu KiSS.2; die Qualifizierungen wurden genannt; zur Frage nach einer Vorgabe eines Prozentsatzes heißt es: „Die Sprachexperten hatten diesmal keine Vorgabe erhalten, einen bestimmten Prozentsatz der Kinder als auffällig zu beurteilen, kennen allerdings (mindestens durch die Zusammenarbeit mit uns) die zu erwartenden Prävalenzen für Sprachentwicklungsstörungen.“
	Fragen nach näheren Angaben zur Testdurchführung durch Studenten	Von den Studenten wurde sowohl der Referenz- als auch der Indextest erhoben; theoretisch konnten die Studenten die Ergebnisse des jeweiligen KiSS.2- oder Referenztests durch ihre Testdurchführung oder Bewertung beeinflussen, nicht aber die endgültigen Untersuchungsergebnisse, d. h. den Vergleich von Index- und Referenztest; um die korrekte Testdurchführung zu prüfen, wurden die Studenten ausführlich angeleitet und kontrolliert
	Nachfrage zur y-Achsen-Beschriftung in Abbildung 1 in der aktuellen Studie [4]	Beschreibung der Bedeutung dieser Beschriftung: Die Achsenbeschriftung „% der Unterstichprobe“ bezieht sich auf den prozentualen Teil aller deutschsprachigen bzw. aller nicht deutschsprachigen Kinder.
Neumann, K. / 18.11.2010	In der E-Mail vom 02.11.2010 wurde von den Autoren erwähnt, dass für die Veröffentlichung der Bewertungsergebnisse seitens des Instituts ggf. eine Genehmigung durch das hessische Sozialministerium benötigt wird; aus diesem Grund wurden die Autoren gebeten, dies mit dem Ministerium zu klären und dem Institut schriftlich mitzuteilen	Neumann, K. / 23.11.2010 bzw. 24.11.2010: Das hessische Sozialministerium wurde um eine schriftliche Genehmigung gebeten.

(Fortsetzung)

Tabelle 3: Dokumentation der Autorenanfragen zur KiSS.2-Studie (Fortsetzung)

Adressat / Datum	Inhalt der Anfrage	Antwort (Antwort durch / Datum / Inhalt)
	Erneute Nachfrage zur Durchführung von Index- und Referenztest durch Studenten: In welcher Reihenfolge wurden Index- und Referenztest durchgeführt? Wurden beim einzelnen Kind Index- und Referenztest von demselben Studenten erhoben?	Die Tests erfolgten nacheinander und in randomisierter Reihenfolge; Index- und Referenztest wurden jeweils von demselben Studenten erhoben
	Nähere Anfrage zur Erfassung der Komorbiditäten	Die Angaben zu Komorbiditäten stammten überwiegend von den Eltern, aber auch von den Erzieherinnen
	Nachfrage nach Zugriffsmöglichkeit der Sprachexpertinnen auf die Ergebnisse des Indextests; Erfragung, ob der Indextest mit oder ohne Kenntnis der Ergebnisse des Referenztests ausgewertet wurde	Die Expertinnen hatten nur Zugriff auf die Ergebnisse des Referenztests; der Indextest wurde von einem Studienmitarbeiter ohne Kenntnis der Ergebnisse des Referenztests ausgewertet
	Detaillierte Anfrage zu der Aussage, dass das im aktuellen Forschungsbericht beschriebene KiSS.2 eine Studienversion war, die in der Nachfolgezeit noch einmal leicht verändert wurde	Begründung und Beschreibung der Veränderung am KiSS.2 in der Nachfolgezeit; die Veränderungen betrafen nicht den Test selbst – also den Screeningbogen, sondern lediglich die an Eltern und Erzieherinnen gerichteten Fragebögen. Demnach enthielt die Studienversion zusätzliche Fragen für Studienzwecke, die im Routineeinsatz nicht von Relevanz sind.
		Neumann, K. / 21.12.2010: Zusendung der Genehmigung des hessischen Sozialministeriums