

Screening auf kritische angeborene Herzfehler mittels Pulsoxymetrie bei Neugeborenen

Vorbericht (vorläufige Nutzenbewertung)

Auftrag: S13-01
Version: 1.0
Stand: 14.11.2014

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Screening auf kritische angeborene Herzfehler mittels Pulsoxymetrie bei Neugeborenen

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

26.06.2013

Interne Auftragsnummer:

S13-01

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Im Mediapark 8 (KölnTurm)

50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Nr. 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Bei dem vorliegenden Vorbericht handelt es sich um eine vorläufige Nutzenbewertung. Zu diesem Vorbericht können Stellungnahmen abgegeben werden, die zu einer Ergänzung und / oder Überarbeitung des Berichts führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG (www.iqwig.de), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vii
Abbildungsverzeichnis	ix
Abkürzungsverzeichnis	x
Kurzfassung	xii
1 Hintergrund	1
2 Ziel der Untersuchung	4
3 Projektbearbeitung	5
3.1 Zeitlicher Verlauf des Projekts	5
3.2 Dokumentation der Änderungen im Projektverlauf	6
4 Methoden	7
4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien in die Nutzenbewertung	9
4.1.1 Population.....	9
4.1.2 Prüf- und Vergleichsintervention	9
4.1.3 Patientenrelevante Endpunkte	9
4.1.4 Studientypen	10
4.1.5 Studiendauer	10
4.1.6 Tabellarische Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien)	10
4.2 Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung	11
4.2.1 Population.....	11
4.2.2 Indextest I und Indextest II.....	11
4.2.3 Referenzstandard	12
4.2.4 Zielgrößen	12
4.2.5 Studientypen	12
4.2.6 Studiendauer	13
4.2.7 Tabellarische Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)	13
4.3 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	13
4.4 Informationsbeschaffung	14
4.4.1 Bibliografische Literaturrecherche	14

4.4.2	Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien	14
4.4.2.1	Systematische Übersichten	14
4.4.2.2	Öffentlich zugängliche Studienregister	14
4.4.2.3	Durch den G-BA übermittelte Unterlagen.....	15
4.4.2.4	Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen	15
4.4.2.5	Informationen aus der Anhörung.....	15
4.4.3	Selektion relevanter Studien.....	15
4.5	Informationsbewertung.....	16
4.5.1	Bewertung von vergleichenden Interventionsstudien	16
4.5.2	Bewertung von Studien zur diagnostischen Güte.....	17
4.6	Informationssynthese und -analyse	17
4.6.1	Gegenüberstellung der Ergebnisse der Einzelstudien	17
4.6.2	Meta-Analysen	18
4.6.2.1	Meta-Analysen für vergleichende Interventionsstudien.....	18
4.6.2.2	Meta-Analysen für Studien zur diagnostischen Güte	19
4.6.3	Sensitivitätsanalysen	20
4.6.4	Subgruppenmerkmale und andere Effektmodifikatoren	20
4.7	Änderungen der Methodik.....	21
5	Ergebnisse.....	23
5.1	Ergebnisse der Informationsbeschaffung	23
5.1.1	Bibliografische Literaturrecherche.....	23
5.1.2	Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien	24
5.1.2.1	Systematische Übersichten	24
5.1.2.2	Öffentlich zugängliche Studienregister	25
5.1.2.3	Durch den G-BA übermittelte Unterlagen.....	25
5.1.2.4	Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen	26
5.1.2.5	Informationen aus der Anhörung.....	27
5.1.3	Resultierender Studienpool	28
5.2	Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudie.....	29
5.2.1	Studiendesign und Studienpopulation	29
5.2.2	Einschätzung des Verzerrungspotenzials auf Studienebene	33
5.3	Ergebnisse zu patientenrelevanten Endpunkten.....	34
5.3.1	Verzerrungspotenzial auf Endpunktebene.....	34
5.3.2	Ergebnisse zum patientenrelevanten Endpunkt kAHF-spezifische Morbidität ...	34
5.3.3	Meta-Analysen	35

5.3.4	Sensitivitätsanalysen	35
5.3.5	Subgruppenmerkmale und andere Effektmodifikatoren	36
5.4	Charakteristika der in die Bewertung eingeschlossenen Studien zur diagnostischen Güte	36
5.4.1	Definitionen von kAHF in den eingeschlossenen Studien und Anpassung an den Bericht	36
5.4.2	Studiendesign und Studienpopulationen	40
5.4.3	Einschätzung des Verzerrungspotenzials	50
5.4.3.1	Verzerrungspotenzial nach QUADAS 2.....	50
5.4.3.2	Bedenken bezüglich der Übertragbarkeit nach QUADAS 2	51
5.5	Ergebnisse zu den Zielgrößen	51
5.5.1	Meta-Analysen	59
5.5.2	Sensitivitätsanalysen	59
5.5.3	Subgruppenmerkmale und andere Effektmodifikatoren	59
5.6	Zusammenfassung der Beleglage.....	59
6	Diskussion.....	61
7	Fazit.....	68
8	Liste der eingeschlossenen Studien	69
9	Literatur	70
Anhang A	– Suchstrategien	76
Anhang B	– Liste der ausgeschlossenen Dokumente mit Ausschlussgründen	80
Anhang C	– Liste der gesichteten systematischen Übersichten	83
Anhang D	– Projektspezifische Vorgaben zur Anwendung von QUADAS 2.....	84
Anhang E	– Relevante Primärstudien und ihr Einschluss in systematische Übersichten (nach ausgewerteten Studien, tabellarische Übersicht)	87
Anhang F	– Leitlinien und Empfehlungen (tabellarische Übersicht).....	88
Anhang G	– Screeningstrategien der eingeschlossenen Studien.....	90

Tabellenverzeichnis

	Seite
Tabelle 1: Kritische angeborene Herzfehler (Übersicht)	1
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien).....	11
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte).....	13
Tabelle 4: In Studienregistern identifizierte relevante Studien.....	25
Tabelle 5: In Studienregistern identifizierte Studien unklarer Relevanz	25
Tabelle 6: Übersicht zur Autorenanfrage	27
Tabelle 7: Studienpool der Nutzenbewertung.....	28
Tabelle 8: Allgemeine Studiencharakteristika	29
Tabelle 9: Übersicht des in der Studie beschriebenen Vergleichs	30
Tabelle 10: Charakterisierung der in der Studie eingeschlossenen Interventionen	32
Tabelle 11: Charakterisierung der Studienpopulation.....	32
Tabelle 12: Verzerrungspotenzial auf Studienebene.....	33
Tabelle 13: Verzerrungspotenzial pro Endpunkt	34
Tabelle 14: Ergebnisse zum Endpunkt: kAHF-spezifische Morbidität (schwere Azidose bei Diagnose).....	35
Tabelle 15: Kritische angeborene Herzfehler: Definitionen, Diagnosen, abweichende Diagnosen (Studiendesign A).....	38
Tabelle 16: kritische angeborene Herzfehler: Definitionen, Diagnosen, abweichende Diagnosen (Studiendesign B).....	39
Tabelle 17: Allgemeine Studiencharakteristika (Studiendesign A)	42
Tabelle 18: Index- und Referenztest (Studiendesign A)	42
Tabelle 19: Eingeschlossene Patienten (Studiendesign A)	43
Tabelle 20: Allgemeine Studiencharakteristika (Studiendesign B)	46
Tabelle 21: Index- und Referenztest (Studiendesign B)	47
Tabelle 22: Eingeschlossene Patienten (Studiendesign B)	48
Tabelle 23: Verzerrungspotenzial QUADAS 2 (Studiendesign A)	50
Tabelle 24: Verzerrungspotenzial QUADAS 2 (Studiendesign B).....	51
Tabelle 25: Bedenken bezüglich der Übertragbarkeit QUADAS 2 (Studiendesign A).....	51
Tabelle 26: Bedenken bezüglich der Übertragbarkeit QUADAS 2 (Studiendesign B).....	51
Tabelle 27: Anzahl zusätzlich gefundener Neugeborener mit kAHF (Studiendesign A)	53
Tabelle 28: Ergebnisse (Studiendesign A)	54
Tabelle 29: Anzahl zusätzlich gefundener Neugeborener mit kAHF (Studiendesign B)	56
Tabelle 30: Ergebnisse (Studiendesign B)	57
Tabelle 31: Relevante Primärstudien und ihr Einschluss in systematische Übersichten	87

Tabelle 32: Screeningempfehlungen ausgewählter Leitlinien	88
Tabelle 33: Internationale Screeningempfehlungen ohne Leitliniencharakter	89
Tabelle 34: Übersicht Pulsoxymetrie-Screeningstrategien	90

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der bibliografischen Literaturrecherche und des Literaturscreenings	24
Abbildung 2: Übersicht über die relevanten Studien aus den verschiedenen Quellen der Informationsbeschaffung und daraus resultierender Studienpool (Studien können in mehreren Quellen identifiziert worden sein).....	28
Abbildung 3: Screeningstrategie de-Wahl Granelli 2009	91
Abbildung 4: Screeningstrategie Koppel 2003	92
Abbildung 5: Screeningstrategie Riede 2010.....	93
Abbildung 6: Screeningstrategie Ruangritnamchai 2007.....	94
Abbildung 7: Screeningstrategie Tautz 2010	95
Abbildung 8: Screeningstrategie Zhao 2014.....	96

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AAP	American Academy of Pediatrics
AHF	angeborener Herzfehler
AS	valvular aortic stenosis (valvuläre Aortenstenose)
CCT	controlled clinical trial (kontrollierte klinische Studie)
s/cCHD	serious/critical congenital heart disease
CoA	coarctation of aorta (Aortenisthmusstenose)
CONSORT	Consolidated Standards of Reporting Trials
D-TGA	dextro- / complete transposition of the great arteries (dextro- / komplette Transposition der großen Arterien)
G-BA	Gemeinsamer Bundesausschuss
HLHS	hypoplastic left heart syndrome (hypoplastisches Linksherzsyndrom)
HTA	health technology assessment
IAA	interrupted aortic arch (unterbrochener Aortenbogen)
ICD	International Classification of Diseases and Related Health Problems (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme)
ICH	International Conference of Harmonization
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITS	intention to screen
ITT	intention to treat
kAHF	kritischer angeborener Herzfehler
NICU	Neonatal Intensive Care Unit
NHS	National Health Service
NNS	Number Needed to Screen
PA / IVS	pulmonary atresia with intact ventricular septum (Pulmonalatresie mit intaktem Ventrikelseptum)
PA / VSD	pulmonary atresia with ventricular septal defect (Pulmonalatresie mit Ventrikelseptumdefekt)
PPV	positive predictive value (positiver prädiktiver Wert)
PS	valvular pulmonary stenosis (Pulmonalstenose)
RCT	randomized controlled trial (randomisierte kontrollierte Studie)
STARD	Standards for the Reporting of Diagnostic Accuracy Studies
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology

Abkürzung	Bedeutung
TAPVC/-D/-R	total anomalous pulmonary venous connection / drainage / return (totale Lungenvenen-Fehlmündung)
TREND	Transparent Reporting of Evaluations with Nonrandomized Designs
U1 / U2	erste / zweite Untersuchung zur Früherkennung von Krankheiten bei Kindern gemäß den Kinder-Richtlinien des G-BA
HHS	US Secretary of Health and Human Services
UVH	univentricular heart (univentrikuläres Herz)
VSD	ventricular septal defect (Ventrikelseptumdefekt)
VOPT-Design	verification of only positive testers

Kurzfassung

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 26.06.2013 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung eines „Screenings auf schwere congenitale Herzfehler mittels Pulsoxymetrie“ beauftragt.

Fragestellung

Ziel der vorliegenden Untersuchung ist die Nutzenbewertung des Pulsoxymetrie-Screenings als Ergänzung (Add-on) zum bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) im Vergleich zum bisherigen diagnostischen Standard ohne Pulsoxymetrie-Screening. Auch das Pulsoxymetrie-Screening als Add-on im Vergleich mit dem bisherigen diagnostischen Standard mit opportunistischer Anwendung der Pulsoxymetrie soll untersucht werden.

Methoden

Es wurden vergleichende Interventionsstudien eingeschlossen, die das Pulsoxymetrie-Screening als Ergänzung (Add-on) der bisherigen Screenings (U1 und U2) oder vergleichbarer klinischer Untersuchungen, gegebenenfalls mit einer opportunistischen Anwendung der Pulsoxymetrie, im Hinblick auf

- Mortalität,
- Morbidität (zum Beispiel schwere Herzinsuffizienz, schwere Hypoxie, kardiogener Schock, schwer ausgeprägter Lungenhochdruck),
- gesundheitsbezogene Lebensqualität des Kindes (gemessen z. B. durch Proxy-Rating),
- psychosoziale Entwicklung (z. B. kommunikative Fähigkeiten, soziale Integration, Selbstkonzeptentwicklung),
- emotionale Entwicklung (z. B. Verhaltensauffälligkeiten),
- grob- und feinmotorische Entwicklung,
- kognitive und bildungsrelevante Entwicklung (z. B. schulisches Leistungsvermögen, Kindergarten / Schulplatzierung / Beschulungsmodus, Ausbildungsmöglichkeiten),
- stationäre Behandlung jeglicher Ursache (Anzahl von Operationen; Verweildauer) und
- unerwünschte Ereignisse

untersuchten. Daneben wurden Studien zur diagnostischen Güte des Pulsoxymetrie-Screenings als Add-on-Test im Vergleich zum bisherigen diagnostischen Standard ohne Pulsoxymetrie mit einer Mindestnachbeobachtungsdauer (bei Nachbeobachtung als Referenztest) von 6 Monaten eingeschlossen.

Hierzu wurde eine systematische Literaturrecherche in den folgenden Datenbanken durchgeführt: MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (Clinical

Trials). Außerdem erfolgte eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanten Primärstudien sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments). Die Suche fand am 15.10.2013 statt.

Darüber hinaus wurden systematische Übersichten, öffentlich zugängliche Studienregister, durch den G-BA übermittelte Unterlagen und die aus dem Anhörungsverfahren zum vorläufigen Berichtsplan zur Verfügung gestellten Publikationen gesichtet. Zudem wurden Autoren von Publikationen relevanter Studien zur Klärung wesentlicher Fragen angeschrieben.

Die Selektion relevanter Studien wurde für das Ergebnis aus der bibliografischen Literaturrecherche, öffentlich zugänglichen Studienregistern und durch den G-BA übermittelten Unterlagen von 2 Reviewern unabhängig voneinander durchgeführt. Die Selektion relevanter Studien aus den übrigen Suchquellen wurde von einem Reviewer durchgeführt und durch einen zweiten überprüft.

Die Datenextraktion erfolgte in standardisierte Tabellen. Zur Einschätzung der Ergebnis-sicherheit wurde das Verzerrungspotenzial auf Studienebene und bei Interventionsstudien zusätzlich auf Endpunktebene bewertet und jeweils in niedrig oder hoch eingestuft; bei Studien zur diagnostischen Güte kam nach QUADAS 2 die Kategorie unklar hinzu. Die Ergebnisse der vergleichenden Interventionsstudie wurden nach Endpunkten geordnet beschrieben. Bei Vergleichbarkeit der Studien hinsichtlich der Fragestellung und relevanter Charakteristika sollten die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden.

Ergebnisse

Insgesamt wurden 6 Studien als relevant für die Fragestellung der vorliegenden Nutzenbewertung identifiziert.

Die Studien untersuchten den Nutzen des Pulsoxymetrie-Screenings als eine den bisherigen diagnostischen Standard ergänzende Untersuchung bei der Erkennung kritischer angeborener Herzfehler zum einen als vergleichende Interventionsstudie (1 Studie) und zum anderen als Studien zur diagnostischen Güte (5 Studien). Studien zur diagnostischen Güte wurden in die Nutzenbewertung einbezogen, weil dem Bericht die Annahme zugrunde gelegt wurde, dass aufgrund der besonderen Konstellation der deutlich positiven Behandlungseffekte von kAHF eine frühere versus eine spätere Diagnose und Therapie mit einer günstigeren Prognose einhergehen.

Das Ergebnis der vergleichenden Interventionsstudie unterstützt die Hypothese, dass ein die klinische Routineuntersuchung ergänzendes Pulsoxymetrie-Screening die krankheits-

spezifische Morbidität bei Neugeborenen mit kAHF verringert (Endpunkt schwere präoperative Azidose (OR: 0,268 [0,110; 0,654], $p = 0,003$). Ein Zusatznutzen des Pulsoxymetrie-Screenings kann aus dieser Studie alleine nicht abgeleitet werden, was im Wesentlichen durch die geringe qualitative Ergebnissicherheit, die unterschiedliche Prävalenz in den Studiengruppen und die inkomplette Nachbeobachtung der gegen Studienende eingeschlossenen Neugeborenen begründet ist.

Vor diesem Hintergrund sind die Ergebnisse der 5 eingeschlossenen Studien zur diagnostischen Güte zu sehen. Aus ihnen ergibt sich, dass mit dem Pulsoxymetrie-Screening zusätzliche Neugeborene mit kAHF entdeckt wurden, die in der klinischen Routineuntersuchung zunächst unauffällig waren.

Mangels ausreichender Datenbasis konnten die Ergebnisse der Studien nicht gepoolt werden, so dass auf die Darstellung der Ergebnisse der einzelnen Studien zurückzugreifen war. In allen Studien wurden mit dem Pulsoxymetrie-Screening zusätzliche Kinder mit kAHF identifiziert. In 2 Studien lagen auch Informationen zur Prävalenz vor. Diese Studien zeigen, dass zwischen 60 und 78 % der Kinder mit kAHF durch das Pulsoxymetrie-Screening zusätzlich entdeckt werden. Bezogen auf die untersuchten Neugeborenen bedeutet dies eine NNS zwischen 421 und 7100: Um 1 zusätzliches Neugeborenes mit kAHF zu finden, mussten in den Studien jeweils zwischen 421 und 7100 asymptomatische Neugeborene ergänzend zur klinischen Routineuntersuchung pulsoxymetrisch gescreent werden.

Der PPV lag zwischen 25,9 % und 75 %: Ein Viertel bis drei Viertel der im Pulsoxymetrie-Screening auffälligen Neugeborenen hatten tatsächlich einen kAHF. Aus dem PPV folgt umgekehrt, dass ebenfalls ein Viertel bis drei Viertel der im Pulsoxymetrie-Screening auffälligen Neugeborenen keinen kAHF hatten. In diesen Fällen falsch-positiver Befunde sind unnötige Therapien möglich. Die Studiendaten zeigen, dass bezogen auf kAHF falsch-positive Befunde in 70 bis 100 % der Fälle auf andere Erkrankungen zurückzuführen und somit nicht intendierte Befunde waren. Das heißt, dass das Pulsoxymetrie-Screening für je 1 mit kAHF identifiziertes Neugeborenes zwischen 2 und 8 Neugeborene identifiziert, die solche nicht intendierten Befunde aufweisen und die nunmehr bereits im asymptomatischen Stadium weiterer Diagnostik und Therapie zugeführt werden. Die Vor- und Nachteile einer Entdeckung und Behandlung dieser Pathologien bereits im asymptomatischen Stadium blieben in den eingeschlossenen Studien insgesamt unklar.

Fazit

Das Pulsoxymetrie-Screening als ein den bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) ergänzendes Screening auf kAHF bei asymptomatischen Neugeborenen zeigt einen Anhaltspunkt für einen Nutzen. Dieses Ergebnis ist auf 1 vergleichende Interventionsstudie und 5 Studien zur diagnostischen Güte zurückzuführen: Die 5 Studien zur diagnostischen Güte zeigen übereinstimmend, dass mit dem Pulsoxymetrie-Screening zusätzliche Neugeborene mit kAHF entdeckt wurden, die in der klinischen Routineuntersuchung unauffällig waren. Das Ergebnis der vergleichenden

Interventionsstudie unterstützt die Hypothese, dass die krankheitsspezifische Morbidität bei Neugeborenen mit kAHF durch das zusätzliche Pulsoxymetrie-Screening verringert werden kann. Der Endpunkt schwere präoperative Azidose bei Neugeborenen mit kAHF zeigte einen statistisch signifikanten Unterschied zugunsten des Pulsoxymetrie-Screenings. Unsicherheit besteht hinsichtlich einer Nutzen-Schaden-Abwägung für die nicht intendierten Befunde, die mit der Pulsoxymetrie zusätzlich erkannt werden.

Schlagwörter: Oxymetrie, Herzfehler – Kongenitale, Nutzenbewertung, Systematische Übersicht

Keywords: Oximetry, Heart Defects – Congenital, Benefit Assessment, Systematic Review

1 Hintergrund

In Deutschland werden von 10 000 Kindern im Mittel 107,6, also etwa 1,1 %, mit einem Herzfehler geboren [1]. Von einem angeborenen Herzfehler (AHF) spricht man bei angeborener Fehlbildung des Herzens oder einzelner Herz- und herznaher Gefäßstrukturen. Das Spektrum von AHF ist breit gefächert. In der Internationalen statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD-10) [2] sind – nicht abschließend – 47 AHF aufgeführt. Eine Systematik von AHF kann anhand unterschiedlicher Kriterien erfolgen. So kann etwa zwischen primär zyanotischen und nicht primär zyanotischen AHF [3] oder zwischen ductusabhängigen und nicht ductusabhängigen AHF [4] differenziert werden. Eine weitere, verbreitete Einteilung von Herzfehlern stellt auf den Schweregrad der AHF ab. Etwa 12 % aller AHF gelten als schwere AHF [1], ca. 10 % werden als kritische (kAHF) bezeichnet [4].

Im Sinne des vorliegenden Berichts sind kAHF solche AHF, bei denen die Fehlbildung des kardiovaskulären Systems derart ausgeprägt ist, dass sie zwar eine weitgehend normale fetale Entwicklung erlaubt, nach der Geburt jedoch eine mit dem (Über-)Leben nicht vereinbare Unterbrechung bzw. Fehlstrukturierung des Blutkreislaufs resultiert. Der Blutkreislauf kann beispielsweise nur durch Kurzschluss- oder Shuntverbindungen aufrechterhalten werden, die eine Umgehung der Hindernisse beziehungsweise der unzureichend ausgebildeten Strukturen oder überhaupt erst eine Verbindung der ansonsten getrennten Kreisläufe ermöglicht (vgl. [4]). In Tabelle 1 sind die kritischen AHF im Sinne des vorliegenden Berichts aufgelistet (Darstellung in Anlehnung an [4]).

Tabelle 1: Kritische angeborene Herzfehler (Übersicht)

Ductusabhängiger Systemkreislauf	unterbrochener Aortenbogen	IAA
	kritische Aortenisthmusstenose	CoA
	kritische Aortenklappenstenose (= valvuläre Aortenstenose)	AS
	hypoplastisches Linksherzsyndrom	HLHS
Ductusabhängiger Pulmonalkreislauf	Pulmonalatresie, Varianten	PA / VSD; PA / IVS
	kritische Pulmonalstenose	PS
	komplexe Herzfehler mit kritischer pulmonaler Minderdurchblutung	
Sonstige und komplexe zyanotische Vitien	totale Lungenvenen-Fehlmündung (mit Obstruktion)	TAPVC/-D/-R
	dextro- / komplette Transposition der großen Arterien	D-TGA
	univentrikuläres Herz, Varianten	UVH

Häufig treten mehrere Herzfehler kombiniert auf; der Schweregrad der Krankheit hängt dann, aber auch bei Vorliegen eines einzigen Herzfehlers, oftmals von der konkreten Ausprägung des kAHF ab (z. B. valvuläre Aortenstenose, Aortenisthmusstenose). Abhängig von der Art des kAHF kann dessen erste Manifestation zu einer rasch progredienten Herzinsuffizienz, zu einer tiefen Hypoxie oder zum kardiogenen Schock und damit zum Tod führen [4]. Ziel einer frühen Intervention ist es, derartige Notfälle zu verhindern, die Hämodynamik zu optimieren und Organschäden zu vermeiden [5]. In der Regel ist sofort nach Diagnose eine medikamentös-stabilisierende Therapie angezeigt. Bei den ductusabhängigen kAHF besteht sie in der Regel in einer Prostaglandin-Behandlung zum Offenhalten des Ductus und damit zur Sicherung des Überlebens der Neugeborenen (vgl. [3]).

2 in der Praxis angewandte Screeningverfahren sind grundsätzlich geeignet, kAHF zu entdecken, ohne dabei gezielte Screeningverfahren auf kAHF zu sein. Im Rahmen des Ultraschall-Screenings gemäß den Mutterschafts-Richtlinien des G-BA im zweiten Schwangerschaftstrimenon können viele kAHF grundsätzlich bereits pränatal erkannt werden. In der am 01.07.2013 in Kraft getretenen Fassung der Mutterschafts-Richtlinien ist in diesem Kontext explizit die Darstellung des Vier-Kammer-Blicks als Teil eines für die schwangere Frau optionalen weiterführenden Ultraschall-Screenings gefordert [6]. Mit dieser Untersuchung wird etwa ein Drittel aller schweren Herzfehler pränatal erkannt [7,8]. Als postnatales Screening erfolgt die klinische Untersuchung. In Deutschland ist diese in der ersten und zweiten Untersuchung zur Früherkennung von Krankheiten bei Kindern (U1 inklusive APGAR-Test und U2) gemäß den Kinder-Richtlinien des G-BA insbesondere mit der Auskultation des Herzens und Tasten der Femoralispulse verortet [9]. Goldstandard zur Überprüfung und Sicherung der Diagnose eines Herzfehlers ist die Echokardiografie, die eine detaillierte Darstellung der Anatomie und Funktion des Herzens sowie die Quantifizierung der Ventrikelfunktion erlaubt.

Pulsoxymetrie ist ein nicht invasives Verfahren zur transkutanen Messung der arteriellen Sauerstoffsättigung (SpO_2) und der Pulsfrequenz. Messort sind der rechte Handballen bzw. ein Finger der rechten Hand (präductal) und / oder der Fuß bzw. eine Zehe des Säuglings (postductal). Ein Fotosensor misst die Absorption von Rotlicht durch das mit Sauerstoff gesättigte Hämoglobin des pulsierend durchströmenden Blutes, das von 2 in einem definierten (Infra-)Rot-Bereich leuchtenden Lichtquellen ausgestrahlt wird. Gleichzeitig wird der Puls in den Kapillaren erfasst. Gemessen wird die Absorption des Lichts in 2 verschiedenen Wellenlängen. Anhand eines Vergleichs des Messergebnisses mit einer Referenztabelle ermittelt ein Überwachungsmonitor den prozentualen Anteil des gesättigten Hämoglobins [10]. Der Trennwert für eine normale Sauerstoffsättigung wird in der Regel bei $< 95\%$ oder $< 96\%$ festgelegt [11-16]. Auch bei primär azyanotischen Vitien wie der valvulären Aortenstenose oder der Aortenisthmusstenose kann bei noch offenem Ductus mit Rechts-Links-Shunt eine in diesem Sinne erniedrigte Sauerstoffsättigung am Fuß gemessen werden.

Ziel eines die bisherigen Screenings ergänzenden, somit zusätzlichen Pulsoxymetrie-Screenings ist es, über den Wert der Sauerstoffsättigung des Blutes einen Hinweis auf das

Vorliegen eines kAHF bei asymptomatischen Neugeborenen zu erhalten und damit die diagnostische Lücke zu verkleinern [13,17-20], die sich darin zeigt, dass ca. 20 bis 25 % der kAHF bis zur klinischen Manifestation unerkannt bleiben [20,21]. Ein Hinweis auf einen kAHF kann mittels der anschließenden abklärenden Echokardiografie entweder verifiziert und das Neugeborene unmittelbar einer Therapie zugeführt werden, oder aber der Hinweis auf einen kAHF wird verworfen. Trifft der Hinweis auf einen kAHF nach Pulsoxymetrie-Screening nicht zu und liegt insofern ein falsch-positives Ergebnis vor, ist zu berücksichtigen, dass eine mit dem Pulsoxymetrie-Screening festgestellte erniedrigte Sauerstoffsättigung häufig Zeichen anderer Erkrankungen ist, wie zum Beispiel neonatale Sepsis, Lungenerkrankungen, persistierende pulmonale Hypertension, andere nicht kritische AHF oder eine Anpassungsstörung [17,19,20,22].

2 Ziel der Untersuchung

Ziel der vorliegenden Untersuchung ist die Nutzenbewertung des Pulsoxymetrie-Screenings als Add-on-Test zum bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) im Vergleich zum bisherigen diagnostischen Standard ohne Pulsoxymetrie-Screening. Auch das Pulsoxymetrie-Screening als Add-on im Vergleich mit dem bisherigen diagnostischen Standard mit opportunistischer Anwendung der Pulsoxymetrie soll untersucht werden.

3 Projektbearbeitung

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 26.06.2013 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung eines „Screenings auf schwere congenitale Herzfehler mittels Pulsoxymetrie“ beauftragt.

Der Auftrag geht auf einen Antrag der Patientenvertretung nach § 140f SGB V vom 06.09.2012 zurück. In diesem Antrag wird unter dem Titel „Screening auf schwere congenitale Herzfehler mittels Pulsoxymetrie nach § 26 SGB V“ ausschließlich der Begriff kritische angeborene Herzfehler verwendet. Erläuternd wird die Definition für kAHF aus der Publikation von Riede und Kollegen 2009 [4] herangezogen. Im Rahmen der Sitzung der Arbeitsgruppe Kinder-Richtlinien des G-BA vom 13.06.2013 in Berlin wurde vor diesem Hintergrund klargestellt, dass die Verwendung des Begriffs „schwer“ auf keine konkrete, insofern auch auf keine von kAHF abweichende Definition abzielt. Aus diesem Grund wird im vorliegenden Bericht die Bezeichnung kritische AHF wie in Kapitel 1 definiert verwendet.

In die Bearbeitung des Projekts sind externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 05.08.2013 Patientenvertreter der Organisationen „Bundesverband Herzkranker Kinder e. V.“, „Elterninitiative herzkranker Kinder Köln e. V.“ und „Herzkind e. V.“ zur Festlegung patientenrelevanter Zielgrößen konsultiert.

Der vorläufige Berichtsplan in der Version 1.0 vom 18.10.2013 wurde am 28.10.2013 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 26.11.2013 konnten schriftliche Stellungnahmen eingereicht werden. Die Dokumentation und Würdigung der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht.

Im Anschluss an die Anhörung wurde ein überarbeiteter Berichtsplan (Version 1.0 vom 03.02.2014) publiziert.

Der vorliegende Vorbericht wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf der Website des IQWiG in einem entsprechenden Leitfaden dargelegt sind. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und/oder Ergänzungen des Berichts führen. Im Anschluss an diese Anhörung wird der Abschlussbericht erstellt. Dieser Bericht wird an den G-BA übermittelt und 8 Wochen später auf der Website des IQWiG veröffentlicht.

3.2 Dokumentation der Änderungen im Projektverlauf

Berichtsplan im Vergleich zum vorläufigen Berichtsplan

Im Vergleich zum vorläufigen Berichtsplan wurde im Berichtsplan auf Nachbeobachtungsdauer von mindestens 1 Jahr bei vergleichenden Interventionsstudien verzichtet (siehe Abschnitt 4.1.5), bei Studien zur diagnostischen Güte wurde sie auf 6 Monate verkürzt (siehe Abschnitt 4.2.6).

Ferner wurden die Zielgrößen präzisiert: Auch Studien mit inkomplettem Referenzstandard (VOPT-Design) waren zu berücksichtigen (siehe Abschnitt 4.2.5).

Redaktionelle Änderungen sind nicht erfolgt.

Zu den methodischen Änderungen siehe ausführlicher Abschnitt 4.7.

Vorbericht im Vergleich zum Berichtsplan

Im Vergleich zum Berichtsplan ergaben sich im Vorbericht lediglich redaktionelle Änderungen.

4 Methoden

Im Folgenden werden die Methoden zur Nutzenbewertung des Pulsoxymetrie-Screenings als Ergänzung (Add-on) der bisherigen Screenings (U1 und U2) oder vergleichbarer klinischer Untersuchungen, gegebenenfalls mit einer opportunistischen Anwendung der Pulsoxymetrie, beschrieben.

Nutzenbewertung anhand von vergleichenden Interventionsstudien (der Screeningkette)

In die Nutzenbewertung wurden prospektiv geplante vergleichende Interventionsstudien einbezogen. Anhand solcher Studien der gesamten Screeningkette mit einer (idealerweise randomisierten) Zuteilung von Personen zu einer Strategie mit beziehungsweise ohne Anwendung der Screeningmaßnahme (gegebenenfalls auch zu unterschiedlichen Screeningstrategien) und der Betrachtung von patientenrelevanten Endpunkten lässt sich der Nutzen von Screeningmaßnahmen bewerten [23]. Zum Beispiel kann zur Nutzenbewertung eines Add-on-Tests ein sogenanntes Strategiedesign angewendet werden [24]. In einer solchen Studie erhält nach Randomisierung der in die Studie eingeschlossenen Neugeborenen die Kontrollgruppe die Versorgung nach bisherigem diagnostischem Standard ohne Pulsoxymetrie-Screening (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) und wird entsprechend dem klinischen Untersuchungsergebnis gegebenenfalls weiterer diagnostischer Abklärung (Echokardiografie) und Therapie zugewiesen. Die Experimentalgruppe erhält ein zusätzliches Pulsoxymetrie-Screening und wird einer aufgrund des Pulsoxymetrie-Screening-Befunds adaptierten, gegebenenfalls abklärenden Diagnostik und gegebenenfalls der entsprechenden Therapie zugeführt.

Da die Pulsoxymetrie als Neugeborenen-Screening auf kAHF in Deutschland bereits in 2 S2-Leitlinien empfohlen wird [14,15] und damit von der Existenz eines sogenannten grauen Screenings auszugehen ist, sollten auch Studien berücksichtigt werden, in denen in der Vergleichsgruppe eine opportunistische Anwendung der Pulsoxymetrie erfolgt.

Nutzenbewertung anhand von Studien zur Bewertung der diagnostischen Güte

Im Rahmen des vorliegenden Berichts wurden auch Studien zur diagnostischen Güte zur Nutzenbewertung herangezogen. Dieses Vorgehen ist möglich, weil bezüglich der Therapie von kAHF deutliche positive Effekte vorliegen und es plausibel erscheint, dass eine frühere Diagnose und Therapie mit einer günstigeren Prognose einhergehen. Dies soll im Folgenden näher erläutert werden:

kAHF im Sinne des vorliegenden Berichts führen definitionsgemäß innerhalb einer kurzen Zeit mit hoher Wahrscheinlichkeit zum Tod, weil es sich um solche Fehlbildungen des kardiovaskulären Systems handelt, die mit dem Überleben nach der Geburt längerfristig nicht vereinbar sind (siehe Kapitel 1). Die Überlebensdauer bei unbehandelten kAHF liegt je nach Art und Ausprägung des kAHF bei wenigen Tagen bis maximal einem Jahr [19]. Für die ductusabhängigen kAHF ist für die Überlebensdauer insbesondere maßgeblich, wann sich der Ductus verschließt [3], was in der Regel innerhalb der ersten Lebensstage geschieht.

Bei den meisten Betroffenen kann durch eine Intervention die Mortalität deutlich verringert werden [25]. Die Ergebnisse der Behandlung von kAHF haben sich durch Fortschritte auf dem Gebiet der Kardiochirurgie und der perioperativen Intensivmedizin in den letzten Jahrzehnten erheblich verbessert. Für nahezu jeden kritischen Herzfehler kann inzwischen eine operative und / oder katheterinterventionelle Therapie angeboten werden, die das Überleben und die Lebensqualität von Kindern mit kAHF deutlich verbessert hat [4,5,25-28]. Das bestätigen Daten zur 1-Jahres-Überlebensrate von Kindern mit kAHF, die in den vergangenen Jahrzehnten insgesamt beständig gestiegen ist [21,29]. Auch das Langzeitüberleben (5 Jahre) von Kindern, die im Zeitraum zwischen 1996 und 2001 mit hypoplastischem Linksherz-Syndrom (HLHS) geboren wurden, war mit 72 % [95 %-KI: 61 %; 83 %] statistisch signifikant höher gegenüber der Gruppe von Kindern, die zwischen 1992 und 1996 geboren wurden (41 % [95 %-KI: 20 %; 51 %]); bei Kindern mit D-TGA hat sich die Überlebensrate aufgrund einer neuen Operationstechnik ebenfalls wesentlich verbessert [30,31].

Aufgrund des plausiblen Wirkmechanismus einer früheren Diagnose und Behandlung ist nicht damit zu rechnen, dass zu dieser Fragestellung (Zusatznutzen einer früheren versus einer späteren Behandlung eines kAHF) prospektiv geplante vergleichende Interventionsstudien durchgeführt werden, wenn hierfür in der Kontrollgruppe diagnostizierte Neugeborene mit kAHF zunächst unbehandelt bleiben müssten. Aufgrund der kurzfristig eintretenden Folgen der kAHF (z. B. beim Schließen des Ductus, siehe Kapitel 1) ist ein derartiges Design nicht durchführbar. Dass solche Studien tatsächlich nicht existieren, bestätigte bereits ein Health-Technology-Assessment-Bericht (HTA-Bericht) des National Health Service (NHS) aus 2005 [32]. Daher bleiben zur Beantwortung dieser Frage nur Beobachtungsstudien mit den entsprechenden methodischen Schwierigkeiten. Eine Zusammenfassung dieser Studien liefert der HTA-Bericht des NHS. Er zeigt, dass eine verspätete Diagnose (hier: Diagnose nach erster Manifestation des kAHF) mit erhöhter Mortalität und Morbidität im Vergleich zu einer früheren Diagnose und Therapie (vor erster Manifestation des kAHF) verbunden ist [32]. Patientenrelevante Nachteile, die sich aus einer früheren gegenüber einer späteren Intervention ergeben, sind hingegen nicht bekannt (vgl. [4]). Etwaige Nachteile, die zum Beispiel mit der prophylaktischen Gabe von Prostaglandin einhergehen könnten, wiegen nicht so schwer, dass sie die zu erwartenden Vorteile – lebensrettende Verhinderung einer rasch progredienten Herzinsuffizienz, einer tiefen Hypoxie oder eines kardiogenen Schocks – infrage stellen könnten.

In Übereinstimmung mit der einschlägigen Literatur wird daher im vorliegenden Bericht vorausgesetzt, dass bei feststehender Diagnose eines kAHF eine frühe gegenüber einer späteren therapeutischen Intervention mit einer besseren Prognose einhergeht (vgl. z. B. [4,13,19,21,22,32-34]).

Problematisch erscheint dagegen, welche Konsequenzen falsch-positive und falsch-negative Ergebnisse des Pulsoxymetrie-Screenings für die Neugeborenen haben. Zu den falsch-positiven Ergebnissen des Pulsoxymetrie-Screenings auf kAHF zählt auch die Entdeckung anderer neonataler Erkrankungen beziehungsweise Auffälligkeiten, die nicht unter die

Definition der kAHF fallen (siehe Kapitel 1). Falsch-positive Befunde (und gegebenenfalls unnötige Therapie) sind in diesen Fällen möglich und müssen daher im Rahmen der Nutzenbewertung mit besonderem Augenmerk bedacht werden. Ebenfalls berücksichtigt werden die möglichen Konsequenzen falsch-negativer Ergebnisse.

Vor dem Hintergrund der in diesem Abschnitt erläuterten Konstellation erfolgte die Nutzenbewertung somit anhand von vergleichenden Interventionsstudien der Screeningkette einerseits und anhand von Studien zur diagnostischen Güte andererseits.

Die Kriterien für den Einschluss von Studien in die Nutzenbewertung werden im Folgenden für vergleichende Interventionsstudien und für Studien zur Bewertung der diagnostischen Güte getrennt dargestellt.

4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien in die Nutzenbewertung

4.1.1 Population

Einzuschließen waren asymptotische Neugeborene.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention hatte aus einem Pulsoxymetrie-Screening als Add-on zum bisherigen Screening im Rahmen der U1- und U2-Untersuchungen oder zu vergleichbaren klinischen Untersuchungen zu bestehen.

Als Vergleichsintervention war das bisherige Screening im Rahmen der U1- und U2-Untersuchungen ohne Pulsoxymetrie-Screening oder vergleichbare klinische Untersuchungen ohne Pulsoxymetrie-Screening gefordert, jeweils mit oder ohne opportunistische Anwendung der Pulsoxymetrie.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung wurden folgende patientenrelevante Endpunkte verwendet:

- Mortalität,
- Morbidität (zum Beispiel schwere Herzinsuffizienz, schwere Hypoxie, kardiogener Schock, schwer ausgeprägter Lungenhochdruck),
- gesundheitsbezogene Lebensqualität des Kindes (gemessen z. B. durch Proxy-Rating),
- psychosoziale Entwicklung (z. B. kommunikative Fähigkeiten, soziale Integration, Selbstkonzeptentwicklung),
- emotionale Entwicklung (z. B. Verhaltensauffälligkeiten),
- grob- und feinmotorische Entwicklung,

- kognitive und bildungsrelevante Entwicklung (z. B. schulisches Leistungsvermögen, Kindergarten / Schulplatzierung / Beschulungsmodus, Ausbildungsmöglichkeiten),
- stationäre Behandlung jeglicher Ursache (Anzahl von Operationen; Verweildauer),
- unerwünschte Ereignisse.

Subjektive Endpunkte (z. B. gesundheitsbezogene Lebensqualität) sollten dann berücksichtigt werden, wenn sie mit validen Messinstrumenten (z. B. validierten Skalen) erfasst wurden.

4.1.4 Studientypen

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für die in Abschnitt 4.1.2 genannten Interventionen und die in Abschnitt 4.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar. Für den zu erstellenden Bericht sollten daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Für den Fall, dass keine einschlägigen RCTs identifiziert werden können, sollten eindeutig prospektiv geplante, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen werden (sogenannte controlled clinical trials [CCTs]), um den patientenrelevanten Zusatznutzen und Schaden des Pulsoxymetrie-Screenings mit ausreichender Ergebnissicherheit schätzen zu können. Dabei war zu beachten, dass in den Studien das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der Studien berücksichtigt wurde. Ferner wurden prospektiv vergleichende Kohortenstudien einbezogen. Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber gegebenenfalls bestärken.

4.1.5 Studiendauer

Für vergleichende Interventionsstudien wurde keine Mindeststudiendauer festgelegt.

4.1.6 Tabellarische Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien)

Tabelle 2 zeigt die Kriterien für den Einschluss von Studien in die Bewertung.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien)

Einschlusskriterien	
E1a	asymptomatische Neugeborene (siehe auch Abschnitt 4.1.1)
E2a	Prüfintervention: Pulsoxymetrie-Screening als Add-on zum bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) (siehe auch Abschnitt 4.1.2)
E3a	Vergleichsintervention: bisheriges Screening im Rahmen der U1- und U2-Screenings ohne Pulsoxymetrie-Screening oder vergleichbare klinische Untersuchungen ohne Pulsoxymetrie-Screening, jeweils mit oder ohne opportunistische Anwendung der Pulsoxymetrie (siehe auch Abschnitt 4.1.2)
E4a	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
E5a	RCTs, CCTs und prospektiv vergleichende Kohortenstudien (siehe auch Abschnitt 4.1.4)
E6a	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [35] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [36], TREND- [37] oder STROBE-Statements [38] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>RCT: randomisierte kontrollierte Studie; CCT: kontrollierte klinische Studie; CONSORT: Consolidated Standards of Reporting Trials; ICH: International Conference of Harmonization; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

4.2 Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung

Aufgrund der besonderen Konstellation der deutlich positiven Behandlungseffekte bei kAHF und der Plausibilität der Annahme, dass eine frühere versus eine spätere Diagnostik und Intervention mit einer besseren Prognose einhergehen (siehe Kapitel 4), wurden im Rahmen des vorliegenden Berichts auch Studien zur diagnostischen Güte zur Nutzenbewertung herangezogen. Charakteristika und Einschlusskriterien für diese Studien werden im Folgenden beschrieben.

4.2.1 Population

Einzuschließen waren asymptomatische Neugeborene.

4.2.2 Indextest I und Indextest II

Als Indextest I wurde das Pulsoxymetrie-Screening als Add-on zum bisherigen Screening im Rahmen der U1- und U2-Untersuchung oder vergleichbare klinische Untersuchungen gefordert.

Als Indextest II wurde das bisherige Screening im Rahmen der U1- und U2-Untersuchungen ohne Pulsoxymetrie-Screening oder vergleichbare klinische Untersuchungen ohne Pulsoxymetrie-Screening akzeptiert, jeweils mit oder ohne opportunistische Anwendung der Pulsoxymetrie.

4.2.3 Referenzstandard

Referenzstandard musste die Echokardiografie oder die Autopsie sowie bei unauffälligem Befund die Nachbeobachtung sein.

4.2.4 Zielgrößen

Eingeschlossen wurden Primärstudien, aus denen personenbezogene Daten zur Berechnung der diagnostischen Güte im Hinblick auf die Entdeckung von kAHF ableitbar sind.

4.2.5 Studientypen

Um die diagnostische Güte des Indextests I zur Erkennung von kAHF bei asymptomatischen Neugeborenen möglichst unverzerrt bestimmen zu können, sollte eine Gruppe von Neugeborenen, die zu einem bestimmten Zeitpunkt prospektiv rekrutiert und mit der Pulsoxymetrie gescreent wurde, zeitnah (Querschnittsdesign) mit dem Referenzstandard (nach-)untersucht werden. Als Referenzstandard für unauffällige Befunde wurde auch die Nachbeobachtung akzeptiert. Notwendig waren ein konsekutiver Einschluss der Neugeborenen und die Dokumentation der fehlenden Werte.

In die Auswertung wurden vergleichende Testgütestudien einbezogen, die den direkten Vergleich des Indextests I zum Indextest II ziehen. Nicht vergleichende Testgütestudien, die lediglich einen indirekten Vergleich zwischen verschiedenen Studien erlauben, ermöglichen aufgrund unvermeidbarer Confounder nur deutlich weniger sichere Aussagen [39].

Studien ohne direkten Vergleich wurden daher nur dann berücksichtigt, wenn nur die nach vorgeschaltetem Indextest II negativen Neugeborenen ein Pulsoxymetrie-Screening erhalten und im Anschluss daran ein Referenztest durchgeführt wurde, während die Indextest-II-positiven Neugeborenen unmittelbar dem Referenztest zugeführt wurden. Nicht berücksichtigt wurden solche Studien, in denen nach vorgeschaltetem Pulsoxymetrie-Screening die Anwendung des Indextests II nur bei den Testnegativen erfolgt (sogenannte Triage-Testanwendung).

Darüber hinaus wurden Primärstudien im VOPT-Design herangezogen, in welchem alle Ergebnisse, die in mindestens einem der beiden Indextests positiv sind, mit dem gleichen, validen Referenzstandard (siehe Abschnitt 4.2.3) untersucht werden [40]. Das VOPT-Design ist eine Abwandlung der Diskordanzstudien, die lediglich die diskordanten Befunde zwischen alternativen diagnostischen Vorgehensweisen untersuchen (inkomplette Verifikation).

4.2.6 Studiendauer

Testnegative Neugeborene sollten mindestens 6 Monate nachbeobachtet werden.

4.2.7 Tabellarische Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)

Die folgende Tabelle 3 zeigt die Kriterien für den Einschluss von Studien in die Bewertung.

Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)

Einschlusskriterien	
E1b	asymptomatische Neugeborene (siehe auch Abschnitt 4.2.1)
E2b	Indextest I: Pulsoxymetrie-Screening als Add-on zum bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) Indextest II: bisheriges Screening im Rahmen der U1- und U2-Untersuchungen ohne Pulsoxymetrie-Screening oder vergleichbare klinische Untersuchungen ohne Pulsoxymetrie-Screening, jeweils mit oder ohne opportunistische Anwendung der Pulsoxymetrie (siehe auch Abschnitt 4.2.2)
E3b	Referenztest: Echokardiografie, Autopsie, gegebenenfalls Nachbeobachtung (siehe auch Abschnitt 4.2.3)
E4b	Zielgrößen: personenbezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt 4.2.4 und 4.2.5)
E5b	direkt vergleichende Diagnosestudien oder nur Pulsoxymetrie-Screening in der Gruppe der Indextest-II-Negativen (siehe auch Abschnitt 4.2.5)
E6b	Mindestnachbeobachtungsdauer 6 Monate bei Nachbeobachtung als Referenzstandard (siehe auch Abschnitt 4.2.6)
E7b	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über Studien, der den Kriterien des STARD-Statements [41] oder des STROBE-Statements [38] genügt. STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology</p>	

4.3 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für das Einschlusskriterium E1a und E1b (Population) war es ausreichend, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt war. Lagen für solche Studien entsprechende Subgruppenanalysen vor, wurde auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1a und E1b bei weniger als 80 % erfüllt war, wurden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorlagen.

Ebenfalls eingeschlossen wurden Studien, die zu mindestens 80 % das Einschlusskriterium E2a (Prüfintervention, bezogen auf die Interventionsgruppe der Nutzenstudie) beziehungsweise das Einschlusskriterium E2b (Indextest I bzw. Indextest II bei Diagnosestudien) erfüllen und bei der Bewertung des patientenrelevanten Nutzens zu mindestens 80 % das Einschlusskriterium E3a erfüllen (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Nutzenstudie).

4.4 Informationsbeschaffung

4.4.1 Bibliografische Literaturrecherche

Die systematische Literaturrecherche nach relevanten Studien wurde in folgenden bibliografischen Datenbanken durchgeführt:

- Suche nach Primärstudien in den Datenbanken MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (Clinical Trials)
- Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanter Primärliteratur sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments)

Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Suche fand am 15.10.2013 statt.

4.4.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien zu ermitteln, wurden weitere Quellen berücksichtigt. Die Rechercheergebnisse wurden anschließend auf weitere relevante Studien und Studienunterlagen untersucht (siehe Abschnitt 4.4.3 „Selektion relevanter Studien“).

4.4.2.1 Systematische Übersichten

Relevante systematische Übersichten wurden hinsichtlich weiterer relevanter Publikationen bzw. Studien gesichtet.

4.4.2.2 Öffentlich zugängliche Studienregister

Die folgenden öffentlich zugänglichen Studienregister wurden durchsucht:

- World Health Organization. International Clinical Trials Registry Platform Search Portal [online]. URL: <http://apps.who.int/trialsearch>

Die Suche in öffentlich zugänglichen Studienregistern fand am 15.10.2013 statt.

4.4.2.3 Durch den G-BA übermittelte Unterlagen

Die vom G-BA mit Auftragserteilung an das IQWiG weitergeleiteten Referenzen wurden hinsichtlich weiterer relevanter Publikationen bzw. Studien gesichtet.

4.4.2.4 Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen

Es war geplant, Anfragen an die Autoren der Publikationen zu stellen, wenn Informationen, die einen relevanten Einfluss auf die Bewertung erwarten ließen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen waren und keine Sponsoren, i. d. R. die Hersteller einer der geprüften Therapieoptionen, zu einer Studie identifizierbar waren.

4.4.2.5 Informationen aus der Anhörung

Im Anschluss an die Veröffentlichung des vorläufigen Berichtsplans erfolgte eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen konnte. Relevante Informationen aus dieser Anhörung konnten in die Nutzenbewertung einfließen.

4.4.3 Selektion relevanter Studien

Selektion relevanter Publikationen aus den Ergebnissen der bibliografischen Literaturrecherche

Die durch die Suche in bibliografischen Datenbanken identifizierten Zitate wurden in einem ersten Schritt anhand ihres Titels und, sofern vorhanden, Abstracts auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 2 und Tabelle 3) bewertet. Als potenziell relevant erachtete Publikationen wurden in einem zweiten Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgten durch 2 Reviewer unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen den beiden Reviewern aufgelöst.

Selektion relevanter Studien aus weiteren Suchquellen

Informationen aus den folgenden Suchquellen wurden von 2 Reviewern unabhängig voneinander auf ihre Relevanz bewertet:

- öffentlich zugängliche Studienregister,
- durch den G-BA übermittelte Unterlagen.

Informationen aus den folgenden Suchquellen wurden von einem Reviewer auf Studien gesichtet, der diese dann auf ihre Relevanz bewertet; ein zweiter Reviewer überprüfte den gesamten Prozess inklusive der Bewertungen:

- im Rahmen der Anhörung zum vorläufigen Berichtsplan eingereichte Informationen.

Die identifizierten relevanten systematischen Übersichten wurden nach weiteren potenziell relevanten Studien durchsucht, deren Relevanz von 2 Reviewern unabhängig voneinander geprüft wurde.

Sofern in einem der genannten Selektionsschritte Diskrepanzen auftraten, wurden diese jeweils durch Diskussion zwischen den beiden Reviewern aufgelöst.

4.5 Informationsbewertung

Die Bewertung der Informationen der eingeschlossenen Studien hängt stark von den verfügbaren Angaben und der Qualität der jeweiligen Publikationen und weiterer Informationsquellen ab. Alle für die Nutzenbewertung relevanten Ergebnisse wurden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft.

4.5.1 Bewertung von vergleichenden Interventionsstudien

Datenextraktion

Alle für die Nutzenbewertung notwendigen Informationen wurden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Besonderes Augenmerk wurde auf die Abbildung der unterschiedlichen Screeningstrategien gerichtet.

Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wurde für jede in die Nutzenbewertung eingeschlossene Studie bewertet, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu wurden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet:

A: Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (bei randomisierten Studien)
- Verdeckung der Gruppenzuteilung (bei randomisierten Studien)
- zeitliche Parallelität der Gruppen (bei nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen beziehungsweise Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung des Patienten sowie der behandelnden Person (bei randomisierten Studien)
- ergebnisgesteuerte Berichterstattung

B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisgesteuerte Berichterstattung

Das Verzerrungspotenzial wurde als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts wurde zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter (A) aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft. Wenn diese Einstufung als „hoch“ erfolgt, wird das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet. Ansonsten finden die unter (B) genannten endpunktspezifischen Aspekte Berücksichtigung.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Nutzenbewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

4.5.2 Bewertung von Studien zur diagnostischen Güte

Es war geplant, die Studien nach Indikationen, Studientypen und gegebenenfalls Strategievergleichen aufzuteilen und separat darzustellen. Der Pool der eingeschlossenen Studien machte eine getrennte Darstellung lediglich nach Studientypen erforderlich. Die Angaben zu Studienpopulationen, methodischen Aspekten und Ergebnissen einzelner Studien wurden tabellarisch dargestellt. Auf die Abbildung der unterschiedlichen Screeningstrategien wurde besonderes Augenmerk gerichtet.

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der Primärstudien zur diagnostischen Güte erfolgte auf Basis des QUADAS-2-Instruments [42]. Das Verzerrungspotenzial von Primärstudien zur diagnostischen Güte wurde als „niedrig“, „unklar“ oder „hoch“ eingestuft.

Eine Einstufung des Verzerrungspotenzials einer Primärstudie als „hoch“ führt nicht zum Ausschluss aus der Bewertung der diagnostischen Güte. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

4.6 Informationssynthese und -analyse

Die Informationen wurden einer Informationssynthese und -analyse unterzogen. Wenn möglich, wurden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Werkzeuge eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen sollte darüber hinaus in jedem Fall erfolgen.

4.6.1 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten (bei Interventionsstudien) und Zielgrößen (bei Studien zur diagnostischen Güte) wurden im Bericht vergleichend beschrieben.

Für bestimmte Fälle war vorgesehen, dass einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt beziehungsweise nicht in die Nutzenbewertung einbezogen werden. Dies trifft insbesondere zu, wenn viele Patienten nicht in der Auswertung enthalten sind. Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basieren, das heißt, wenn der Anteil der Patienten ohne jegliche Berücksichtigung in der Auswertung (Nichtberücksichtigungsanteil) größer als 30 % ist. In der Literatur werden zum Teil bereits Nichtberücksichtigungsanteile größer als 20 % als nicht mehr aussagekräftig betrachtet [43].

Ausnahmen von dieser Regel können zum Beispiel dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [44].

Die Ergebnisse sollten auch dann nicht in die Nutzenbewertung oder Bewertung der diagnostischen Güte einbezogen werden, wenn der Unterschied der Nichtberücksichtigungsanteile zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.6.2 Meta-Analysen

4.6.2.1 Meta-Analysen für vergleichende Interventionsstudien

Es war geplant, bei hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbaren Studien die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammenzufassen. Für die statistische Auswertung sollten primär die Ergebnisse aus ITT-Analysen verwendet werden, so wie sie in den vorliegenden Dokumenten beschrieben sind.

Meta-Analysen erfolgen in der Regel auf Basis von Modellen mit zufälligen Effekten [45]. In begründeten Ausnahmefällen werden Modelle mit festen Effekten eingesetzt. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, werden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

Für stetige Variablen wird die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g , als Effektmaß eingesetzt. Bei binären Variablen werden Meta-Analysen primär anhand des Odds Ratio und beim Vergleich von Überlebenszeiten anhand des Hazard Ratio durchgeführt. In begründeten Ausnahmefällen kommen auch andere Effektmaße zum Einsatz. Bei kategorialen Variablen wird ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und von den verfügbaren Daten verwendet [46].

Die Effektschätzer und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgt die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität [46]. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ($p \geq 0,2$ für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive

Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität werden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.6.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.6.4).

4.6.2.2 Meta-Analysen für Studien zur diagnostischen Güte

Es war beabsichtigt, aus den Studien die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle mittels Forest Plots [47] zusammenfassend darzustellen. Außerdem sollte für die Testgütekriterien anhand der Sensitivität und Spezifität eine bivariate Meta-Analyse durchgeführt werden, sofern die dafür nötigen Anforderungen erfüllt waren [48].

Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [49,50]. Der Algorithmus zum Schätzen der Parameter im bivariaten Modell kann zu unpräzisen Schätzungen führen, das heißt zu Schätzungen mit zu großen Standardfehlern und entsprechenden Konfidenzregionen. Auch kann der Algorithmus gegebenenfalls keine Schätzungen liefern, wenn das Maximum-Likelihood-Verfahren nicht konvergiert. In beiden Fällen fehlen brauchbare Schätzungen. Die Gründe hierfür können beispielsweise sein, dass zu wenige Studien vorliegen oder dass einzelne Studien extreme Werte aufweisen. Sind die resultierenden Schätzungen unpräzise, werden die Ergebnisse der bivariaten Meta-Analysen in der Regel nicht dargestellt.

Falls die bivariate Meta-Analyse präzise Schätzungen liefert, so werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Ergebnisse verschiedener Strategien, die aus derselben Studie stammen, werden durch eine Verbindungslinie gekennzeichnet. Des Weiteren werden die aus der bivariaten Meta-Analyse gewonnenen Schätzungen für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95 %-Konfidenzregionen dargestellt [51].

Auf einen Unterschied in der diagnostischen Güte zwischen 2 Strategien wird mit dem Likelihood-Ratio-Test (LRT) getestet (Test auf Gleichheit der Erwartungswerte von Sensitivität und Spezifität der beiden Strategien). Die Differenzierung zwischen einer „höheren diagnostischen Güte“ und einer „vermutlich höheren diagnostischen Güte“ wird anhand der grafischen Darstellungen vorgenommen.

Liefert die bivariate Meta-Analyse unpräzise Schätzungen, die zu sehr großen Konfidenz- und Prädiktionsregionen führen, so werden keine Aussagen über einen Unterschied in der diagnostischen Güte zwischen 2 Strategien abgeleitet. In Ausnahmefällen, wie beispielsweise beim Vorliegen von mehreren großen Studien mit niedrigem Verzerrungspotenzial, werden die Ergebnisse geeigneter univariater statistischer Tests, das heißt für die Sensitivität und Spezifität getrennt, dargestellt. Zusammen mit den grafischen Darstellungen werden gege-

benenfalls Aussagen zu einem Unterschied in der diagnostischen Güte zwischen 2 Strategien abgeleitet.

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen untersucht.

4.6.3 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse waren Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant.

Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Trennwerten oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ wird für Sensitivitätsanalysen verwendet.

Das Ergebnis der Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufter Effekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen Nutzen attestiert wird. Bei Studien zur diagnostischen Güte können 95 %-Prädiktionsregionen für den Erwartungswert einer neuen Studie bestimmt werden, um Ausreißer zu identifizieren. Sensitivitätsanalysen nach dem Leave-one-out-Prinzip können durchgeführt werden, um den Einfluss jeder einzelnen Studie auf das Ergebnis zu bewerten.

4.6.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollten hinsichtlich potenzieller Effektmodifikatoren untersucht werden, das heißt hinsichtlich klinischer Faktoren, die die Effekte beeinflussen.

Effektmodifikatoren können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.6.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren.

Es war vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter (in Stunden / Lebenstagen),

- Art und Schweregrad der AHF,
- pränatale Diagnostik im Hinblick auf einen kritischen AHF,
- Messort (prä- und / oder postductal),
- Art der Entbindung.

Weitere mögliche Effektmodifikatoren, die sich aus den verfügbaren Informationen ergeben, sollten ebenfalls begründet einbezogen werden können.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines Zusatznutzens auf eine spezielle Subgruppe von Patienten eingeschränkt werden.

4.7 Änderungen der Methodik

Berichtsplan im Vergleich zum vorläufigen Berichtsplan

Im Zuge des Stellungnahmeverfahrens zum vorläufigen Berichtsplan hatte sich die Erkenntnis durchgesetzt, dass die dort geforderte mindestens einjährige Nachbeobachtung nicht zwingend erforderlich ist. Der Rahmen für den Einschluss sowohl von vergleichenden Interventionsstudien als auch von Studien zur diagnostischen Güte wird dadurch erweitert.

- Verzicht auf die Festlegung einer Mindeststudiendauer bei vergleichenden Interventionsstudien

Bei vergleichenden Interventionsstudien scheint der Verzicht auf die Mindeststudiendauer von einem Jahr bereits aufgrund des Add-on-Charakters des Pulsoxymetrie-Screenings methodisch gerechtfertigt. Maßgeblich kommt es darauf an, ob durch Pulsoxymetrie mehr Neugeborene mit kAHF identifiziert werden können als durch die klinische Untersuchung allein. Eine Schlechterstellung von Neugeborenen, die ein falsch-negatives Pulsoxymetrie-Screening-Testergebnis haben, gegenüber Neugeborenen, die kein Pulsoxymetrie-Screening erhalten haben, ist unwahrscheinlich, da alle Neugeborenen auch ein Testergebnis aus der klinischen Untersuchung erhalten haben. Vor diesem Hintergrund sind die aus einer Nachbeobachtung zu gewinnenden Daten (falsch-negative Testergebnisse) verzichtbar. Aus diesem Grund wurde „E6a Mindestnachbeobachtungsdauer 1 Jahr“ gestrichen.

- Verkürzung der Mindeststudiendauer und Präzisierung der Zielgrößen bei Studien zur diagnostischen Güte

Bei Studien zur diagnostischen Güte wurde die Mindeststudiendauer auf 6 Monate verkürzt. Das trägt dem Umstand Rechnung, dass die im Berichtsplan festgelegten Zielerkrankungen sich in der Regel spätestens innerhalb dieses Zeitraums manifestieren.

Zusätzlich wurden die Zielgrößen dahin gehend präzisiert, dass bei nicht ausreichender Anzahl von Testgütestudien mit vollständigem Referenzstandard auch Studien mit in-

komplettem Referenzstandard (VOPT-Design) berücksichtigt werden können. Da im Rahmen eines derartigen Studiendesigns positive Ergebnisse zeitnah mit dem Referenzstandard (Echokardiografie) verglichen werden und negative Ergebnisse nicht ausgewertet werden, entfällt hier eine Nachbeobachtung gänzlich. So kann dem Umstand besser Rechnung getragen werden, dass beim Screening auf kAHF der Anteil richtig-positiver Testergebnisse wichtiger ist als der Anteil falsch-positiver.

Vorbericht im Vergleich zum Berichtsplan

Im Vergleich zum Berichtsplan ergeben sich im Vorbericht keine Änderungen der Methodik.

5 Ergebnisse

In die Nutzenbewertung des Pulsoxymetrie-Screenings als Add-on-Test zum bisherigen diagnostischen Standard sollten neben vergleichenden Interventionsstudien auch Studien zur diagnostischen Güte einfließen (siehe Kapitel 4). Da die Literatursuche für beide Studientypen mit einer Recherche abgedeckt wurde, sind die Ergebnisse der Informationsbeschaffung zusammenfassend dargestellt. Im Übrigen erfolgt die Darstellung der Ergebnisse zu den beiden Studientypen getrennt.

5.1 Ergebnisse der Informationsbeschaffung

5.1.1 Bibliografische Literaturrecherche

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche nach Studien in den bibliografischen Datenbanken und des Literaturscreenings gemäß den Kriterien zum Studieneinschluss.

Nach Ausschluss von 261 Duplikaten ergab sich eine Gesamtzahl von 565 zu screenenden Treffern.

519 Treffer wurden von beiden Reviewern gegebenenfalls nach Konsentierung zunächst diskrepanter Einschätzungen übereinstimmend im Rahmen des Titel- und Abstractscreenings als nicht relevant ausgeschlossen. Aus der bibliografischen Literaturrecherche verblieben damit 46 potenziell relevante Treffer, die im Volltext gesichtet wurden. Hiervon wurden 35 aufgrund fehlender Relevanz ausgeschlossen. Die Zitate der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B. Bei weiteren 6 Treffern handelte es sich um relevante systematische Übersichten, welche im Hinblick auf relevante Studien gescreent wurden (siehe Abschnitt 5.1.2.1).

Die verbliebenen 5 Publikationen zu 5 Studien erfüllten nach übereinstimmender Einschätzung beider Reviewer die für diesen Bericht definierten Kriterien zum Studieneinschluss. Insgesamt wurden damit aus dieser bibliografischen Recherche 1 Interventionsstudie und 4 Studien zur diagnostischen Güte eingeschlossen.

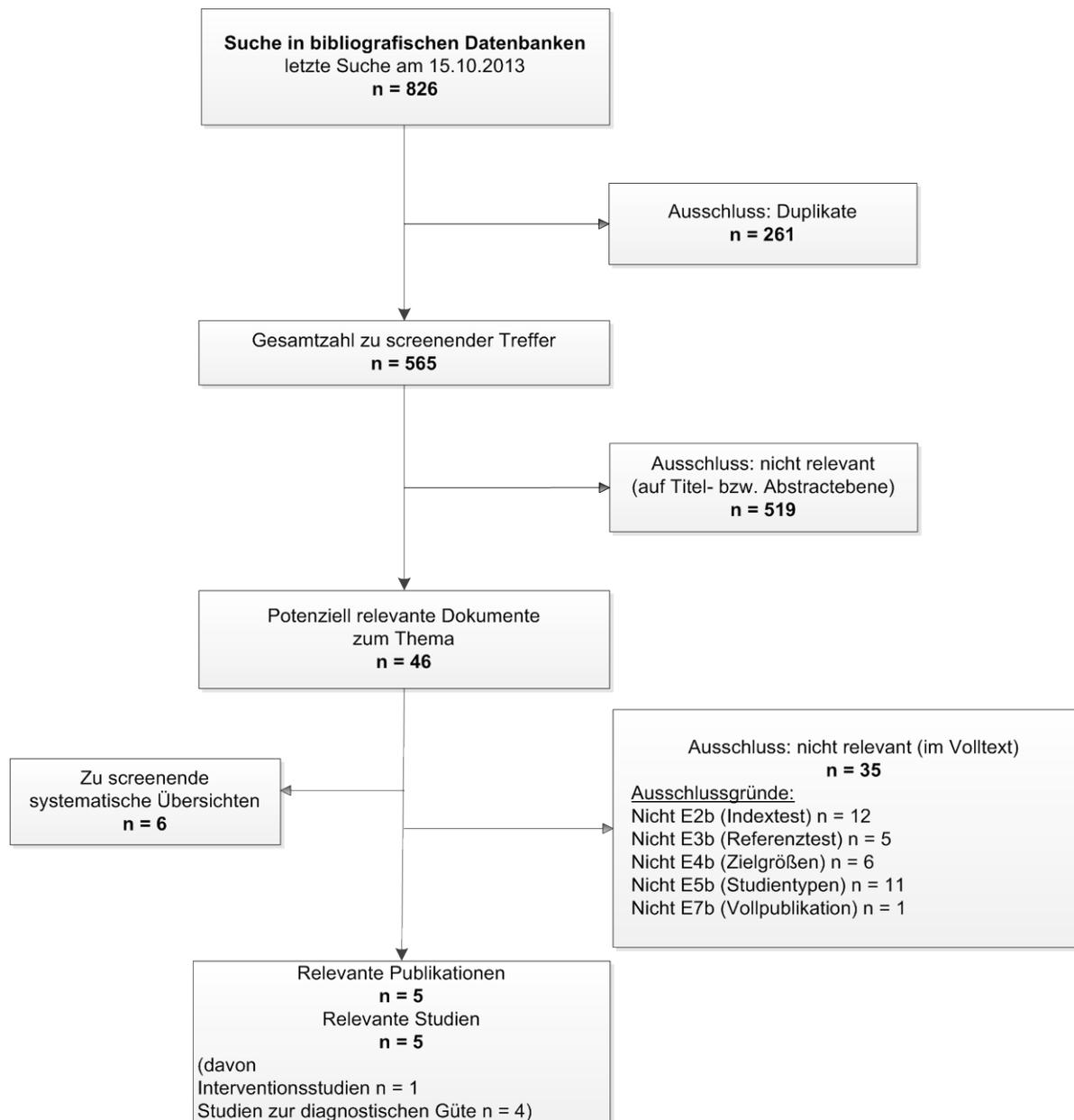


Abbildung 1: Ergebnis der bibliografischen Literaturrecherche und des Literaturscreenings

5.1.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien

5.1.2.1 Systematische Übersichten

Im Rahmen der bibliografischen Literaturrecherche wurden 6 relevante systematische Übersichten identifiziert. Deren Sichtung ergab keine weiteren relevanten Publikationen bzw. Studien, die nicht über andere Rechenschritte identifiziert wurden.

Im Rahmen der Sichtung der mit Auftragserteilung an das IQWiG weitergeleiteten Unterlagen des G-BA wurden keine weiteren relevanten systematischen Übersichten identifiziert.

5.1.2.2 Öffentlich zugängliche Studienregister

Durch die Suche nach weiteren publizierten und nicht publizierten Studien in den in Abschnitt 4.4.2.2 genannten öffentlich zugänglichen Studienregistern wurde die folgende relevante Studie identifiziert:

Tabelle 4: In Studienregistern identifizierte relevante Studien

Studienregister ID	Studie	Studienregister	Status	Ergebnisbericht in Studienregister vorhanden
UKCRN ID: 4134	Pulse Oximetry as a screening test for congenital heart disease in newborn babies (PULSEOX)	UK Clinical Research Network: Portfolio Data	abgeschlossen	nein

Die relevante Studie konnte auch über die bibliografische Literaturrecherche identifiziert werden [18,19].

Darüber hinaus wurde eine Studie unklarer Relevanz identifiziert.

Tabelle 5: In Studienregistern identifizierte Studien unklarer Relevanz

Studienregister ID	Studie	Studienregister	Status	Ergebnisbericht in Studienregister vorhanden
NCT01665261	Newborn Screening for Congenital Heart Disease (NSCHD)	Clinicaltrials.gov	laufend	nein

Für laufende Studien wurden keine Autorenanfragen gestellt.

5.1.2.3 Durch den G-BA übermittelte Unterlagen

Insgesamt wurden 57 Referenzen vom G-BA mit Auftragserteilung an das IQWiG weitergeleitet.

Diese wurden auf Duplikate zur bibliografischen Literaturrecherche und im Rahmen des Titel- und Abstractscreenings auf auszuschließende Dokumente überprüft. Nach übereinstimmender Einschätzung beider Reviewer verblieb damit 1 im Volltext zu sichtendes Dokument [52].

1 Dokument wurde von beiden Reviewern im Rahmen der Volltextsichtung gegebenenfalls nach Konsentierung zunächst diskrepanter Einschätzungen als nicht relevant ausgeschlossen. Die Zitate der ausgeschlossenen Dokumente finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B.

Kein Dokument erfüllte nach übereinstimmender Einschätzung beider Reviewer die für diesen Bericht definierten Kriterien für den Studieneinschluss.

5.1.2.4 Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen

Für 6 Studien wurden Autorenanfragen versendet; darunter auch zu einer weiteren Studie, die mittels Autoalert ermittelt und bereits für die Erstellung des Vorberichts eingeschlossen wurde (Zhao 2014, siehe Abschnitt 5.1.3). Eine Übersicht zu den Autorenanfragen befindet sich in der nachfolgenden Tabelle 6. Die Informationen aus den eingegangenen Antworten sind in die Studienbewertung eingeflossen.

Tabelle 6: Übersicht zur Autorenanfrage

Studie	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
de-Wahl Granelli 2009	<ul style="list-style-type: none"> ▪ Daten zur Vervollständigung der Vierfeldertafel ▪ Klärung differierender Angaben in der Publikation ▪ Anforderung des Studienprotokolls ▪ Grenzwertangabe (schwere Azidose) 	ja	<ul style="list-style-type: none"> ▪ Daten sind für die Autorin zurzeit nicht zugänglich ▪ Klärung erfolgt ▪ Überlassung des Flussdiagramms / Screeningprotokolls ▪ Angabe des Grenzwerts
Koppel 2003	<ul style="list-style-type: none"> ▪ prospektive Festlegung des Trennwerts (Indextext I) ▪ Anforderung des Studienprotokolls 	ja	<ul style="list-style-type: none"> ▪ Trennwert wurde prospektiv festgelegt ▪ Überlassung des Protokolls für die Screeningstrategie
Riede 2010	<ul style="list-style-type: none"> ▪ prospektive Festlegung des Trennwerts (Indextext I) ▪ Angaben zum Referenzstandard (Nachbeobachtung) 	ja	<ul style="list-style-type: none"> ▪ Trennwert wurde prospektiv festgelegt ▪ Vollständigkeit der Nachbeobachtung durch Erfassung der Neugeborenen / Säuglinge mit verspäteter Diagnose
Ruangritnamchai 2007	<ul style="list-style-type: none"> ▪ prospektive Festlegung des Trennwerts (Indextext I) ▪ Anforderung des Studienprotokolls 	nein	k. A.
Tautz 2010	<ul style="list-style-type: none"> ▪ prospektive Festlegung des Trennwerts (Indextext I) ▪ Anforderung des Studienprotokolls 	ja	<ul style="list-style-type: none"> ▪ Trennwert wurde prospektiv festgelegt ▪ Überlassung des Protokolls für die Screeningstrategie
Zhao 2014	<ul style="list-style-type: none"> ▪ Daten zur Vervollständigung der Vierfeldertafel ▪ Klärung inkonsistenter Angaben in der Publikation 	ja	<ul style="list-style-type: none"> ▪ Daten zur Vervollständigung der Vierfeldertafel ▪ Klärung differierender Angaben in der Publikation z. T. durch Korrektur
k. A.: keine Angabe			

5.1.2.5 Informationen aus der Anhörung

Im Rahmen der Anhörung zum vorläufigen Berichtsplan wurden keine zusätzlichen relevanten Studien genannt.

5.1.2.6 Zusätzlich identifizierte relevante Studien

Mittels kontinuierlicher Beobachtung der Literatur (Autoalert) wurde eine weitere relevante Studie [53] identifiziert, die nicht über andere Rechenschritte gefunden werden konnte. Diese Studie wurde nach Abschluss der bibliografischen Literaturrecherche für den Vorbericht publiziert. Sie wurde gleichwohl bereits für die Erstellung des Vorberichts

eingeschlossen. Eine systematische Recherche zur Überprüfung der Aktualität und Vollständigkeit des Studienpools wird im Rahmen des Abschlussberichts erfolgen.

5.1.3 Resultierender Studienpool

Die nachfolgende Abbildung 2 fasst die Ergebnisse der durchsuchten Quellen zusammen. Für jede Quelle wird die Anzahl der dort als relevant identifizierten Studien dargestellt, wie sie auch in den vorangegangenen Abschnitten beschrieben wurden. Zudem wird dargestellt, wie groß der daraus resultierende Studienpool für die Nutzenbewertung ist.

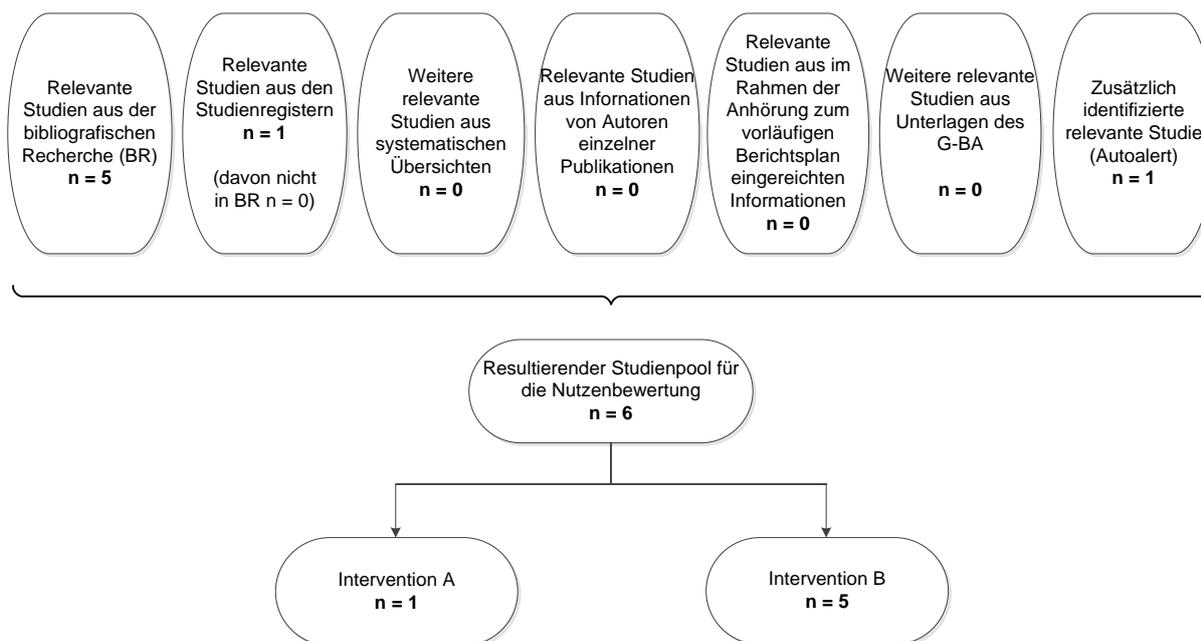


Abbildung 2: Übersicht über die relevanten Studien aus den verschiedenen Quellen der Informationsbeschaffung und daraus resultierender Studienpool (Studien können in mehreren Quellen identifiziert worden sein)

Durch die verschiedenen Suchschritte konnten insgesamt 6 relevante Studien (6 Publikationen) identifiziert werden (siehe auch Tabelle 7).

Tabelle 7: Studienpool der Nutzenbewertung

Studie	Vollpublikation (in öffentlich zugänglichen Fachzeitschriften)
de-Wahl Granelli 2009	ja [17]
Koppel 2003	ja [54]
Riede 2010	ja [20]
Ruangritnamchai 2007	ja [55]
Tautz 2010	ja [56]
Zhao 2014 ^a	ja [53]

a: Nach Abschluss der systematischen Recherche eingeschlossene Studie.

5.2 Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudie

5.2.1 Studiendesign und Studienpopulation

Die Nutzenbewertung schließt 1 vergleichende Interventionsstudie ein: de-Wahl Granelli 2009 (Tabelle 8 bis Tabelle 11). In dieser Studie untersuchte das schwedische Forscherteam den Einfluss des Screenings als Add-on zu einer klinischen Routineuntersuchung auf patientenrelevante Endpunkte und verglich die Ergebnisse mit denen aus einer Kontrollgruppe, in der kein zusätzliches Pulsoxymetrie-Screening durchgeführt wurde (siehe Tabelle 8).

In 5 Zentren in Westgötaland war das Pulsoxymetrie-Screening auf kAHF in die routinemäßige Neugeborenenversorgung integriert. Für die Kontrollgruppe wurden weitere Zentren aus der Umgebung herangezogen, in denen diese zusätzliche Untersuchung nicht Teil der Routine war.

In die Studie wurden alle Neugeborenen eingeschlossen, bei denen nicht bereits vorgeburtlich ein kAHF diagnostiziert worden war (siehe Tabelle 9). Neben asymptomatischen erfüllten damit auch symptomatische Neugeborene die Einschlusskriterien, wobei davon auszugehen ist, dass der Anteil symptomatischer Neugeborener jedenfalls > 20 % lag (s. Abschnitt 4.1.3). de-Wahl Granelli und Kollegen definierten kAHF als „lebensbedrohliche ductusabhängige angeborene Herzfehler“. Diese Definition ist enger als die des vorliegenden Berichts, weil sie die Entdeckung anderer als der ductusabhängigen kAHF als falsch-positive Befunde einordnet, zum Beispiel sonstige und komplexe zyanotische Vitien wie TAPVR, siehe Tabelle 1.

Tabelle 8: Allgemeine Studiencharakteristika

Studie	Studienland	Studiendesign	Studienziel	Patientenrelevante Zielkriterien	Evidenzstufe ^a	Verzerrungspotenzial
de-Wahl Granelli 2009	Schweden	prospektive vergleichende Kohortenstudie	<ul style="list-style-type: none"> ▪ Evaluation des Nutzens des Pulsoxymetrie-Screenings zur frühen Entdeckung von kAHF ▪ Vergleich der Detektionsraten von kAHF durch klinische Untersuchung plus Pulsoxymetrie-Screening und klinische Untersuchung allein 	<ul style="list-style-type: none"> ▪ kAHF-spezifische Morbidität (schwere Azidose bei Diagnose) 	IIb	hoch
a: entspricht der Evidenzstufe des G-BA kAHF: kritische angeborene Herzfehler						

Tabelle 9: Übersicht des in der Studie beschriebenen Vergleichs

Studie	Interventionen	Studiendauer	Patienten (EK / AK)	Zentren (Ort) Zeitraum
de-Wahl Granelli 2009	Prüfintervention: POS + klinische Routineuntersuchung	33 Monate	<ul style="list-style-type: none"> ▪ EK: Neugeborene ▪ AK: pränatal diagnostizierter kAHF 	Westgötaland: <ul style="list-style-type: none"> ▪ Östra Hospital ▪ Mölndal Hospital ▪ Borås Hospital ▪ Skövde Hospital ▪ Trollhättan Hospital 7/2004–3/2007
	Vergleichsintervention: klinische Routineuntersuchung	48 Monate		Krankenhäuser in anderen Regionen 1/2004–12/2007
AK: Ausschlusskriterium; EK: Einschlusskriterium; kAHF: kritische angeborene Herzfehler				

In der Interventionsgruppe in Westgötaland nahmen Hebammen und Krankenschwestern bei 39 821 Neugeborenen im Alter von durchschnittlich 38 Stunden prä- und postductal, also an der rechten Hand und an einem Fuß, den Wert der funktionellen Sauerstoffsättigung (SpO₂). Nach Abzug von 1470 Ausschlüssen verblieben 38 429 Neugeborene (81,8 % der Geburtenkohorte, n = 46 963). Als Trennwert legten de-Wahl Granelli und Kollegen unter Berufung auf eine eigene frühere Studie aus 2005 [57] persistierend < 95 % Sauerstoffsättigung prä- und postductal oder eine Differenz zwischen beiden Messorten von > 3 % fest.

Nach dem Pulsoxymetrie-Screening folgte bei 38 374 Neugeborenen eine verblindete klinische Routineuntersuchung (81,7 % der Geburtenkohorte, n = 46 963). Die untersuchenden Kinderärzte hatten bei diesen Untersuchungen keine Kenntnis vom Ausmaß der Sauerstoffsättigung der Neugeborenen und beurteilten daher ausschließlich aufgrund der klinischen Untersuchung, ob ein schwacher, ein starker oder kein Verdacht auf einen kAHF vorlag. Weitere Einzelheiten der klinischen Routineuntersuchung sind nicht berichtet. Es war vorgesehen, dass die Verblindung dann aufgehoben wird, wenn das Neugeborene einen SpO₂-Wert von < 90 % zeigt. In diesen Fällen (n = 55) sollte das Neugeborene noch am gleichen Tag echokardiografisch untersucht werden. Ansonsten erfolgte die Echokardiografie bei jedem positiv getesteten Neugeborenen, sei es beim Pulsoxymetrie-Screening aufgefallen oder bei der klinischen Untersuchung.

Vollständige Untersuchungsprotokolle sowohl vom Pulsoxymetrie-Screening als auch von der verblindeten klinischen Routineuntersuchung lagen insgesamt von 38 374 Neugeborenen vor.

In der Kontrollgruppe kamen im Untersuchungszeitraum 108 604 Kinder zur Welt; sie alle durchliefen die übliche klinische Routineuntersuchung. Die Untersuchungsinhalte sind nicht explizit berichtet. Es wurde davon ausgegangen, dass die Untersuchung derjenigen der Studiengruppe entsprach und genau wie diese die Kontrolle der Herzfunktion umfasste (siehe auch das Flussdiagramm in Anhang G, Abbildung 3).

Für die Nachbeobachtung griffen die Forscher auf klinische Daten ebenso wie auf öffentliche Register zurück.

In der Interventionsgruppe wurde bei 62 von 46 963 Neugeborenen ein kAHF festgestellt; 2 davon waren bereits pränatal diagnostiziert worden. In der Kontrollgruppe war der Anteil von Neugeborenen mit kAHF etwas geringer: Insgesamt wurden 109 Neugeborene mit kAHF gefunden, davon waren 9 bereits pränatal diagnostiziert (siehe Tabelle 10 und Tabelle 11).

Tabelle 10: Charakterisierung der in der Studie eingeschlossenen Interventionen

Studie	Prüfintervention: Pulsoxymetrie-Screening + klinische Routineuntersuchung Geburtenkohorte: n = 46 963					Vergleichs- intervention: klinische Routineuntersuchung Geburtenkohorte: n = 108 604		Nachbeobachtung
	Pulsoxymetrie-Screening (funktionelle Sauerstoffsättigung) n = 38 429 (81,8 %)			klinische Routineuntersuchung (verblindet) n = 38 374 (81,7 %)		Testzeitpunkt	Ablauf	
	Messort	Testzeitpunkt	Trennwert	Testzeitpunkt	Trennwert			
de-Wahl Granelli 2009	Fuß und rechte Hand	38 Stunden (median age, range 1–406 h)	<ul style="list-style-type: none"> ▪ persistierend (3 Messungen) <ul style="list-style-type: none"> ▫ < 95 % präductal und postductal oder ▫ Unterschied präductal/postductal > 3 % ▪ (initial) ≤ 90 % 	nach POS	kein Verdacht / schwacher Verdacht / starker Verdacht auf kAHF	k. A.	k. A.	<ul style="list-style-type: none"> ▪ klinische Nachbeobachtung ▪ Registeranalyse (nationale Datenbank der nationalen Behörde für forensische Medizin)
k. A.: keine Angabe; kAHF: kritische angeborene Herzfehler								

Tabelle 11: Charakterisierung der Studienpopulation

Studie		Prävalenz kAHF insgesamt	Prävalenz kAHF ohne pränatal diagnostizierte kAHF	Geschlecht	Alter
de-Wahl Granelli 2009	Interventionsgruppe	62/46 963 (0,13 %)	60/46 961 (0,13 %)	k. A.	Neugeborene
	Kontrollgruppe	109/108 604 (0,10 %)	100/108 599 (0,09 %)	k. A.	
k. A.: keine Angabe; kAHF: kritische angeborene Herzfehler					

5.2.2 Einschätzung des Verzerrungspotenzials auf Studienebene

de-Wahl Granelli und Kollegen haben ihre Studie als prospektive vergleichende Interventionsstudie durchgeführt. Nicht randomisierte vergleichende Studien besitzen aufgrund der fehlenden Randomisierung generell ein hohes Verzerrungspotenzial [23]. Bereits aus diesem Grund war das Verzerrungspotenzial der Studie als hoch einzustufen.

Da es sich in Interventions- und Kontrollgruppen um Geburtseinrichtungen mit vergleichbarer Versorgungsstufe handelte, sind Gruppenunterschiede zum Beispiel in Hinblick auf den Anteil symptomatischer Neugeborener, Gestationsalter (insbesondere Frühgeburten), Geburtsart (vaginale Entbindung oder Kaiserschnitt), Alter oder Suchtverhalten der Mutter (Nikotin, Alkohol, Betäubungsmittel etc.) unwahrscheinlich. Ein Einfluss von „Störvariablen“ (Confounder) ist jedoch nicht mit Sicherheit auszuschließen. Die Angabe der sogenannten Baseline-Charakteristika fehlt. Sie hätte solche Unterschiede offengelegt und eine Adjustierung erlaubt. Unklar blieb ferner die Verblindung von Behandlern und Patienten (hier: die Eltern oder Sorgeberechtigten der Neugeborenen) in Bezug auf die Gruppenzuteilung. Unabhängig von der Gruppenzuteilung erfolgte in der Interventionsgruppe die klinische Untersuchung in Unkenntnis der Ergebnisse der pulsoxymetrischen Untersuchung und insofern verblindet (zu Ausnahmen siehe Abschnitt 5.2.1). Unter „Mortalität“ wurde nicht die Gesamtmortalität oder die Letalität unter Kindern mit kAHF berichtet, sondern lediglich die Todesfälle von Kindern mit undiagnostiziertem kAHF. Daher war unklar, ob alle relevanten Endpunkte ergebnisunabhängig berichtet wurden.

Das Verzerrungspotenzial auf Studienebene wurde anhand definierter Indikatoren beurteilt. Tabelle 12 bildet die Ergebnisse der Bewertung ab.

Tabelle 12: Verzerrungspotenzial auf Studienebene

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen / adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung Patienten ^a	Verblindung Behandler ^b	Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte	Fehlen sonstiger (endpunkttüber-greifender) Aspekte, die das Verzerrungspotenzial beeinflussen	Einstufung des Verzerrungspotenzials auf Studienebene
de-Wahl Granelli 2009	ja ^c	unklar ^d	unklar	unklar	unklar ^e	ja	hoch
<p>a: Hier: Eltern anstelle der Patienten. b: Verblindung hinsichtlich der Anwendung des / der Indextests. c: Der Beobachtungszeitraum der Prüfintervention liegt innerhalb des Beobachtungszeitraums der Vergleichsintervention. d: Keine Angaben zu Baseline-Charakteristika, keine Adjustierung für Confounder. e: Angabe der Anzahl von Todesfällen bei Kindern mit undiagnostiziertem kAHF als Mortalität; keine Angabe der Mortalität bei allen Kindern mit kAHF oder Gesamtmortalität.</p>							

5.3 Ergebnisse zu patientenrelevanten Endpunkten

5.3.1 Verzerrungspotenzial auf Endpunktebene

War schon das Verzerrungspotenzial auf Studienebene als „hoch“ einzustufen, so galt das selbe für die Bewertung auf Endpunktebene (siehe Tabelle 13).

Tabelle 13: Verzerrungspotenzial pro Endpunkt

Studie	Endpunkt	Verblindung der Endpunkte-erheber	Adäquate Umsetzung des ITS-Prinzips	Ergebnis-unabhängige Bericht-erstattung dieses Endpunkts allein	Fehlen sonstiger (endpunkt-spezifischer) Aspekte, die das Verzerrungs-potenzial beeinflussen	Einstufung des Verzerrungs-potenzials der Ergebnisse des Endpunkts
de-Wahl Granelli 2009	kAHF-spezifische Morbidität (schwere Azidose bei Diagnose)	unklar ^a	ja ^b	ja	ja	hoch
a: Sofern das Ereignis im Geburtskrankenhaus eintritt: keine Verblindung, sonst: Verblindung (faktisch) wahrscheinlich. b: Eigene ITS-Auswertung anhand der Daten. ITS: intention to screen						

5.3.2 Ergebnisse zum patientenrelevanten Endpunkt kAHF-spezifische Morbidität

Die Ergebnisse zu dem Endpunkt kAHF-spezifische Morbidität finden sich in Tabelle 14.

Die kAHF-spezifische schwere Azidose wurde als Surrogat für unmittelbar eintretende schwerwiegende Beeinträchtigungen akzeptiert. In de-Wahl Granelli 2009 wurde der Grenzwert für eine schwere Azidose bei einem pH-Wert < 7,2 festgelegt (siehe Abschnitt 5.1.2.4 zu Autorenanfragen). Aus Leitlinien und Fachliteratur ergibt sich die Patientenrelevanz einer schweren Azidose, die es unbedingt zu vermeiden gilt. Eine deutsche Leitlinie führt die Azidose mit einem Nabelarterien-pH-Wert von < 7,0 als absolute Indikation für eine Verlegung aus einer Geburtsklinik in eine Klinik mit der adäquaten Versorgungsstufe an [58].

Bei den in den Tabellen berichteten Ergebnissen handelt es sich zunächst um eigene Berechnungen, in denen die Verhältnisse von krankheitsspezifischen Ereignissen zu der Anzahl eingeschlossener Neugeborener in den jeweiligen Gruppen dargestellt sind (Interventionsgruppe: n = 46 961; Kontrollgruppe n = 108 599). Dieses Vorgehen entspricht dem ITS-Prinzip (intention to screen), wonach in die Berechnung die Gesamtheit der zu

screenenden Gruppe einzustellen ist, um den Nutzen und Schaden des Pulsoxymetrie-Screenings für die Gesamtheit der gescreenten Kinder zu beurteilen.

In der Interventionsgruppe hatten 7 Neugeborene mit kAHF eine schwere präoperative Azidose entwickelt, in der Kontrollgruppe waren es 33 Neugeborene. Der Gruppenunterschied zeigt mit einem Odds Ratio von 0,490 (95 %-KI: [0,217; 1,109]) einen p-Wert von 0,086.

In der zweiten Zeile der Tabelle finden sich die von den Autoren berichteten Daten zu krankheitsspezifischen Ereignissen im Verhältnis zur Anzahl von Kindern mit kAHF (Interventionsgruppe: n = 60; Kontrollgruppe n = 100). Auf dieser Grundlage kommen die Studienautoren zu einem statistisch signifikanten Ergebnis für den Endpunkt kAHF-spezifische schwere präoperative Azidose (p = 0,0025).

Tabelle 14: Ergebnisse zum Endpunkt: kAHF-spezifische Morbidität (schwere Azidose bei Diagnose)

Studie	Prüfintervention	Vergleichsintervention	Gruppenunterschied
de-Wahl Graneli 2009	7/46 961 (0,01 %) ^a	33/108 599 (0,03 %) ^a	p = 0,086 ^{a, b} OR 0,490 [0,217; 1,109] ^a
	7/60 (11,67 %)	33/100 (33 %)	P = 0,003 ^b OR: 0,268 [0,110; 0,654] ^a
a: eigene Berechnung b: Exakter Test von Fisher. OR: Odds Ratio			

Über die Ergebnisse zur Morbidität hinaus berichtet de-Wahl Graneli 2009 des Weiteren Daten zur Mortalität. Diese Daten geben Aufschluss über das Risiko, mit undiagnostiziertem kAHF aus dem Krankenhaus entlassen zu werden. Hierfür wurde die Anzahl an Kindern herangezogen, die mit in der Geburtsklinik undiagnostiziertem kAHF verstorben sind. Diese Angaben erlauben keine Rückschlüsse für den patientenrelevanten Endpunkt Mortalität. Hierfür wäre die Angabe der Gesamtmortalität oder der kAHF-spezifischen Mortalität jeweils in beiden Gruppen erforderlich; Daten hierzu wurden jedoch nicht berichtet. Somit lässt sich nicht feststellen, ob in der Interventionsgruppe mehr oder weniger Todesfälle insgesamt oder bei Kindern mit kAHF aufgetreten sind als in der Kontrollgruppe.

5.3.3 Meta-Analysen

Weil nur eine vergleichende Interventionsstudie in die Nutzenbewertung eingeschlossen werden konnte, kam eine Meta-Analyse nicht in Betracht.

5.3.4 Sensitivitätsanalysen

Aufgrund der Datenlage entfielen Sensitivitätsanalysen.

5.3.5 Subgruppenmerkmale und andere Effektmodifikatoren

Subgruppenanalysen wurden in der eingeschlossenen Studie nicht durchgeführt.

5.4 Charakteristika der in die Bewertung eingeschlossenen Studien zur diagnostischen Güte

Für die Bewertung der diagnostischen Güte des Pulsoxymetrie-Screenings ist die Definition von kAHF von ausschlaggebender Bedeutung. Bevor die Charakteristika der in die Nutzenbewertung eingeschlossenen Studien beschrieben werden, ist daher der Umgang mit den uneinheitlichen Definitionen von kAHF in den eingeschlossenen Studien zu erläutern.

5.4.1 Definitionen von kAHF in den eingeschlossenen Studien und Anpassung an den Bericht

Wie weit oder eng der Begriff der kritischen Herzfehler gefasst wird, wirkt sich unmittelbar auf die diagnostische Güte des Pulsoxymetrie-Screenings aus: Bei einem weiten Verständnis von kAHF werden auch solche durch Pulsoxymetrie entdeckten Herzfehler als richtig-positive Untersuchungsergebnisse gewertet, die bei einem engeren Verständnis nicht die Kriterien für einen kAHF erfüllten und daher als falsch-positiv einzuordnen wären. Die Weite der Definition von kAHF hat auch Auswirkungen auf die Annahmen, die dem vorliegenden Bericht zugrunde liegen, dass bezüglich der Therapie von kAHF deutliche positive Effekte vorliegen und es plausibel erscheint, dass eine frühere Diagnose und Therapie mit einer günstigeren Prognose einhergehen. Diese Annahmen wären für schwächer ausgeprägte Herzfehler nicht unbedingt haltbar.

Die Forschergruppen der eingeschlossenen Studien haben, mit einer Ausnahme, ihre jeweils eigene Definition von kAHF. Einen Überblick über die Definitionen und die Einordnung der gefundenen Herzfehler als kAHF im Sinne des Berichtsplans oder andere Diagnosen geben die Tabelle 15 und Tabelle 16.

Um die eingeschlossenen Studien in Hinblick auf ihre Aussagen zur diagnostischen Güte des Pulsoxymetrie-Screenings vergleichen zu können, wurde diesem Bericht eine Definition von kAHF vorangestellt (siehe Kapitel 1), die mit den Definitionen aus den einzelnen Publikationen abgeglichen wurde, und sodann eine entsprechende Kategorisierung der einzelnen Herzfehler vorgenommen.

Zhao 2014 greift die Klassifizierung von Herzfehlern von Ewer und Kollegen [19] auf. In der Kategorie „major congenital heart disease“ wird zwischen „critical congenital heart disease“ (cCHD) und „serious congenital heart disease“ (sCHD) unterschieden. cCHD sind solche Herzfehler, die (unbehandelt) innerhalb von 28 Tagen nach der Geburt entweder zum Tod führen oder interventionsbedürftig sind. In Abgrenzung davon werden AHF, die nach dem ersten Lebensmonat, aber innerhalb des ersten Jahres interventionsbedürftig sind, als sCHD bezeichnet.

Der Studie Koppel 2003 liegt eine Definition von „kritischen angeborenen kardiovaskulären Fehlbildungen“ zugrunde. Umfasst sind Läsionen, die „wahrscheinlich einer chirurgischen Korrektur in den ersten Lebensmonaten“ bedürfen.

Fast in wörtlicher Übereinstimmung sind in Ruangritnamchai 2007 die von ihnen mittels Pulsoxymetrie-Screening gesuchten kAHF definiert: Läsionen, die wahrscheinlich einer chirurgischen Korrektur in den ersten (hier:) wenigen Lebensmonaten bedürfen.

Ähnlich lautet auch die Definition in Riede 2010. Unter kAHF wurde eine „Gruppe morphologisch heterogener Fehlbildungen gefasst, denen gemeinsam ist, dass zur Lebensrettung ein früher chirurgischer Eingriff oder eine Katheterintervention erfolgen muss“.

Bei Tautz 2010 haben die Autoren in ihrer Publikation auf eine Definition von kAHF verzichtet und lediglich alle gefundenen kAHF berichtet.

Vor dem Hintergrund dieser unterschiedlichen (beziehungsweise fehlenden) Definitionen war es erforderlich, die in den Publikationen zu den Studien berichteten kAHF entsprechend der Definition des Berichts neu zu kategorisieren und zwischen kAHF und „anderen Diagnosen“ zu unterscheiden (siehe Tabelle 15 und Tabelle 16).

Aus dieser Neuordnung entsprechend der berichtseigenen Definition ergeben sich teilweise Abweichungen zwischen den von den Forschergruppen berichteten und den in den Ergebnistabellen wiedergegebenen Daten.

Tabelle 15: Kritische angeborene Herzfehler: Definitionen, Diagnosen, abweichende Diagnosen (Studiendesign A)

Studie	Definition (k)AHF der Studie	Anzahl identifizierter Neugeborener gemäß berichtseigener Definition von kAHF	Diagnosen von kAHF gemäß berichtseigener Definition (Anzahl, nach Art der Entdeckung: Pulsoxymetrie-Screening (P), Indextest II (K), Nachbeobachtung (NB))	Andere Diagnosen (Anzahl Neugeborener)
Zhao 2014	„major CHD: defects causing death or requiring intervention during infancy: critical (defects causing death or needing intervention before 28 days of age) and serious (defects needing intervention before 1 year of age) CHD“	222 ^a (davon cCHD: 146 ^b sCHD: 98 ^c)	cCHD: Critical PS ^{P(10), K(10)} (10) ToF ^{P(9), K(9)} (9) TAC ^{P(2), K(3), NB(1)} (5) Single ventricle ^{P(8), K(9), NB(1)} (11) PA ^{P(30), K(28)} (30) TGA ^{P(32), K(29), NB(1)} (33) DORV ^{P(8), K(6)} (9) HLHS ^{P(3), K(2), NB(3)} (7) CoA ^{P(3), K(4), NB(2)} (7) IAA ^{P(2), K(2), NB(1)} (5) Critical AS ^{P(1), K(3)} (3) TAPVC ^{P(14), K(8), NB(1)} (17) sCHD^d: PS (10) ToF (46) TAC (3) Single ventricle (10) PA/VSD (4) DORV (+ complex) (11) Ebstein (6) COA (+ complex) (7) TAPVC (3) AS (2)	VSD (39) PDA (7) AVSD (25) other congenital heart disease (90) persistent pulmonary hypertension (41) lung problem (23) infection (10) preterm birth (16)

a: Weitere 22 kAHF fielen in der (unsystematischen) Nachbeobachtung auf; 10 davon bei den cCHD, die übrigen bei den sCHD (s. Spalte 4; letztere ohne Zuordnung zu den einzelnen Herzfehlern, s. Endnote c).
 b: Im Appendix zur Publikation sind 11 weitere cCHD aufgeführt, die bei symptomatischen Neugeborenen entdeckt wurden und daher hier abzuziehen waren.
 c: Im Appendix zur Publikation sind 4 weitere sCHD aufgeführt, die bei symptomatischen Neugeborenen entdeckt wurden. Anders als bei den cCHD lässt sich nicht zuordnen, um welche Herzfehler es sich dabei handelt. Daher ergibt sich die weitere Diskrepanz zwischen 3. (n = 222) und 4. Spalte (n = 248).
 d: Für sCHD ist die Art der Entdeckung (Pulsoxymetrie, klinische Untersuchung oder Nachbeobachtung) nicht publiziert. Auf Autorenanfrage konnte jedoch die Anzahl pro Indextest ermittelter kAHF entsprechend der berichtseigenen Definition ermittelt werden.
 kAHF: kritischer angeborener Herzfehler;
 CHD: congenital heart disease; cCHD: critical congenital heart disease; sCHD: serious congenital heart disease. AA: aortic atresia, AP: aorto-pulmonary, AS: aortic stenosis, ASD: atrial septal defect, AVSD: atrioventricular septal defect, CoA: coarctation of the aorta, DILV: double inlet left ventricle, DORV: double outlet right ventricle, HLHS: hypoplastic left heart syndrome, IAA: interrupted aortic arch, PA: pulmonary atresia, PS: pulmonary stenosis, SAS: subvalvular aortic stenosis, TA(C): truncus arteriosus communis, TAPVC/R: total anomalous pulmonary venous connection/return, TGA: transposition of the great arteries, ToF: tetralogy of fallot, VSD: ventricular septal defect.

Tabelle 16: kritische angeborene Herzfehler: Definitionen, Diagnosen, abweichende Diagnosen (Studiendesign B)

Studie	Definition (k)AHF der Studie	Anzahl identifizierter Neugeborener gemäß berichtseigener Definition von kAHF	Diagnosen von kAHF gemäß berichtseigener Definition (Anzahl, nach Art der Entdeckung: Pulsoxymetrie-Screening (P), Indextest II (K), Nachbeobachtung (NB))	Andere Diagnosen (Anzahl, nach Art der Entdeckung: Pulsoxymetrie-Screening (P))
Koppel 2003	Critical congenital cardiovascular malformation (critical CCVM): lesion that would likely require surgical correction during the first month of life.	5	TAPVR ^P (2) TAC ^P (1) CoA ^{NB} (1) Hypoplastic left pulmonary artery with aorto-pulmonary collaterals ^{NB} (1)	Persistent right to left ductal shunting ^P (1)
Riede 2010	„[A] group of morphologically heterogeneous disorders which have in common that early surgical or catheter interventional therapy is mandatory to achieve survival.“	18	TA with IAA ^P (1) HLHS ^P (1) PA with VSD ^P (2) TAPVD ^P (5) TGA ^P (2) TGA with VSD ^P (2) Taussig Bing syndrome ^P (1) CoA ^{NB} (3) TGA with VSD ^{NB} (1) kAHF ^K (18) (nicht spezifiziert)	PPHN ^P (15) Sepsis ^P (13)
Ruangritnamchai 2007	„[A] lesion that would likely require surgical correction during the first few months of life.“ ^{ca}	1	TGV ^P (1)	AV canal (= AVSD) & moderate TR ^P (1) VSD, PFO, moderate TR, PI ^P (1)
Tautz 2010	(keine Definition)	8	TGA ^P (3) TAPVR ^P (2) TAC ^P (1) CoA ^P (1) AS + PAPVR ^P (1)	PPHN ^P (2) Infection ^P (7) AV canal (= AVSD) ^P (1)

a: Die Definitionen der Autoren umfassen auch Herzfehler, die nicht als kritische angeborene Herzfehler im hier verstandenen Sinne einzuordnen sind. Daher weichen die Zahlen der hier dargestellten kAHF von denen in den Primärpublikationen ab.

kAHF: kritischer angeborener Herzfehler;

AS: aortic stenosis, AV canal: complete atrioventricular canal (= AVSD), CoA: coarctation of the aorta, HLHS: hypoplastic left heart syndrome, IAA: interrupted aortic arch, LPA: left pulmonary artery, PA: pulmonary atresia, PAPVR: partial anomalous pulmonary venous return, PDA: patent ductus arteriosus, PFO: Patent Foramen Ovale, PI: pulmonary insufficiency, PPHN: Persistent pulmonary hypertension of the newborn, PS: Pulmonary stenosis, TA(C): truncus arteriosus (communis), TAPVD/R: total anomalous pulmonary venous drainage/return, TGA: transposition of the great arteries, TGV: Transposition of the great vessels, TR: tricuspid regurgitation, VSD: ventricular septal defect.

5.4.2 Studiendesign und Studienpopulationen

In die Nutzenbewertung sind die Ergebnisse aus 5 Studien zur diagnostischen Güte eingeflossen (zum Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung siehe Kapitel 4). Auch die Studie de-Wahl Granelli 2009 berichtet Ergebnisse zur diagnostischen Güte. Da in dieser Studie kein direkter Vergleich im Sinne einer Anwendung beider Indextests am gleichen Neugeborenen angestellt wurde, konnte sie nicht in die vorliegende Analyse eingeschlossen werden (zur Interventionsgruppe siehe Abschnitt 5.2.1). Tabelle 17 bis Tabelle 22 bieten einen Überblick über die Studiencharakteristika, unter anderem zum Studiendesign, zu den untersuchten Zielgrößen, den Indextests und Referenztests sowie zur Studienpopulation.

Bei den eingeschlossenen Studien lassen sich 2 Designs unterscheiden. Zhao 2014 untersuchte die diagnostische Güte der Pulsoxymetrie (Indextest I) einerseits und der klinischen Routineuntersuchung (Indextest II) andererseits, indem alle eingeschlossenen Neugeborenen beide Indextests erhielten (Studiendesign A). Die übrigen Studien dagegen wählten ein negatives Testergebnis der klinischen Routineuntersuchung als Einschlusskriterium und damit als Bedingung für die Durchführung des Pulsoxymetrie-Screenings (Studiendesign B).

Der Unterschied besteht darin, dass die zuerst genannte Strategie (Studiendesign A) sowohl Erkenntnisse darüber liefert, ob Neugeborene mit kAHF über das Pulsoxymetrie-Screening *und* gleichermaßen die klinische Routineuntersuchung identifiziert wurden, als auch Erkenntnisse über diskrepante Ergebnisse, wenn also die pulsoxymetrische Untersuchung ein positives Ergebnis lieferte und die klinische Routineuntersuchung ein negatives und umgekehrt.

Im Unterschied dazu treffen Studien nach dem zweiten Studiendesign ausschließlich eine Aussage dazu, wie viele Neugeborene nach negativer klinischer Routineuntersuchung (Indextest II) durch das Pulsoxymetrie-Screening (Indextest I) identifiziert wurden, mithin wie viele zusätzliche Neugeborene mit kAHF das Pulsoxymetrie-Screening im Vergleich zur klinischen Routineuntersuchung gefunden hat – ohne zu prüfen, ob die durch die klinische Untersuchung gefundenen kAHF auch durch das Pulsoxymetrie-Screening gefunden worden wären.

Da die vorgestellten Prüfschemata Ergebnisse mit unterschiedlichem Aussagewert liefern, werden die Charakteristika und Ergebnisse im Folgenden nach den beiden Gruppen „Studiendesign A“ und „Studiendesign B“ sortiert berichtet.

Studiendesign A

In der chinesischen Studie **Zhao 2014** ging es um die Bewertung der diagnostischen Güte des Pulsoxymetrie-Screenings sowie der klinischen Routineuntersuchung einzeln und in Kombination (siehe Tabelle 17). In dieser Studie sollten alle eingeschlossenen Neugeborenen beide Indextests erhalten. Auf eine systematische Nachbeobachtung wurde verzichtet, sodass

die Daten für die negativen Ergebnisse nicht übernommen werden konnten und somit auf die VOPT-Auswertung zurückzugreifen war.

Die funktionelle Sauerstoffsättigung wurde in dieser Studie prä- und postductal gemessen (rechte Hand und ein Fuß, Indextest I). Die Neugeborenen waren im Mittel 43 Stunden alt; frühester Messzeitpunkt waren 6, spätester 72 Stunden nach der Geburt. Der Trennwert wurde unter Bezugnahme auf die Empfehlung des US Secretary of Health and Human Services (HHS) [59] (siehe dazu Kapitel 6, Abschnitt Ergebnisse und Schlussfolgerungen von Leitlinien und Empfehlungen) auf persistierend (2 Messungen) $< 95\%$ prä- und postductaler Sättigung, $> 3\%$ Differenz zwischen prä- und postductaler Sättigung und initial $< 90\%$ Sauerstoffsättigung festgelegt.

Die klinische Untersuchung (Indextest II) erfolgte (zur Sicherstellung der Unkenntnis des klinischen Untersuchers von der Sauerstoffsättigung) unmittelbar vor dem Pulsoxymetrie-Screening. Anhand von 4 definierten Kriterien sollte die Einstufung als testpositiv oder testnegativ erfolgen (siehe Tabelle 18).

Test-positive Neugeborene aus Indextest I und Indextest II waren echokardiografisch nachzuuntersuchen. Negativ getestete Neugeborene aus beiden Indextests wurden (unsystematisch) nachbeobachtet, wobei hier auf eine klinische Nachuntersuchung nach 6 Wochen und Elternfeedback bei eintretendem Ereignis gesetzt wurde (siehe Tabelle 18, siehe auch das Flussdiagramm in Abbildung 8, Anhang G.)

Eingeschlossen wurden alle im Studienzeitraum Neugeborenen mit Ausnahme derjenigen, bei denen bereits die pränatale Diagnose eines kAHF bestand. Auch symptomatische Neugeborene wurden eingeschlossen, die Auswertung erfolgte jedoch getrennt von den asymptomatischen. Von 130 282 für diese Studie geeigneten Neugeborenen erhielten 120 707 asymptomatische Neugeborene beide Indextests, davon waren 55 % männlich. Der Anteil durch diese Tests identifizierter Neugeborener mit kAHF lag bei 222 / 120 707 (0,2 %) (siehe Tabelle 19). Der Anteil pränatal identifizierter Kinder mit kAHF lag mit 8,7 % unter dem Durchschnitt, der für Deutschland angenommen wird (siehe Kapitel 1).

Tabelle 17: Allgemeine Studiencharakteristika (Studiendesign A)

Studie	Studienland	Studiendesign	Studienziel	Zielkriterium	Evidenzstufe ^a
Zhao 2014	China	prospektive Studie	Diagnostische Güte der Kombination aus Pulsoxymetrie-Screening und klinischer Untersuchung speziell auf kAHF ^b	Bewertung der diagnostischen Güte von Pulsoxymetrie-Screening und klinischer Untersuchung auf kAHF einzeln und in Kombination	IIb
<p>a: Entspricht der Evidenzstufe des G-BA. b: Die Studie umfasst 2 Studiengruppen: symptomatische und asymptomatische Neugeborene. In den Ergebnistabellen wird ausschließlich über die Ergebnisse aus der Gruppe der asymptomatischen Neugeborenen berichtet. G-BA: Gemeinsamer Bundesausschuss; kAHF: kritischer angeborener Herzfehler</p>					

Tabelle 18: Index- und Referenztest (Studiendesign A)

Studie	Indextest I: Pulsoxymetrie-Screening			Indextest II: klinische Routineuntersuchung		Referenztest(s)
	Messort	Testzeitpunkt (in Lebensstunden)	Trennwert ^a	Testzeitpunkt (Mittel oder in Relation zu Indextest I)	Trennwert	
Zhao 2014	Fuß und rechte Hand	43 (Mittel; Spanne: 6–72)	<ul style="list-style-type: none"> ▪ ≤ 90 % in der ersten Messung, ▪ 90 %–95 % prä- und postductal in 2 Messungen oder ▪ >3 % Differenz prä- / postductal in 2 Messungen 	vor Pulsoxymetrie-Screening	4 Indikatoren: Familien-Anamnese, spezielle Gesichtszüge, Herzgeräusch, extrakardiale Fehlbildungen	<ul style="list-style-type: none"> ▪ Echokardiografie bei Index-I- und -II-Positiven; ▪ (unsystematische) Nachbeobachtung der Index-I- und -II-Negativen (klinische Nachbeobachtung nach 6 Wochen, Eltern-Feedback bei Ereignis)
a: funktionelle Sauerstoffsättigung						

Tabelle 19: Eingeschlossene Patienten (Studiendesign A)

Studie	Eingeschlossene Neugeborene (n)				Einschlusskriterien	Ausschlusskriterien	Neugeborene mit kAHF	Anteil pränatal identifizierter Neugeborener mit kAHF	Geschlecht % w/m	Alter (in Lebensstunden zum Testzeitpunkt)
	geeignet (n)	Indextest I (n)	Indextest II (n)	Indextests I+II (n)						
Zhao 2014	128 224 ^a	120 707 ^b	120 707	120 707	alle Neugeborenen	pränatale Diagnose eines kAHF	222 ^c	8,7 % (24 ^d /274)	45/55	43 (Durchschnittsalter, Spanne: 6–72)

a: In die Studie wurden 2031 symptomatische und 128 251 asymptomatische Neugeborene eingeschlossen. Das Screening und die Auswertung erfolgten getrennt. Die Angaben in dieser und den folgenden Tabellen beziehen sich ausschließlich auf die Gruppe der asymptomatischen Neugeborenen. In dieser Gruppe wurden Neugeborene mit pränatal diagnostiziertem AHF ausgeschlossen (n = 27), sodass die Anzahl der geeigneten Neugeborenen 128 224 beträgt.

b: darunter Frühgeborene < 37 Gestationswochen: n = 13 695 (< 20 %)

c: Von den 315 in der Publikation als „major“ klassifizierten Herzfehlern waren 71 postnatal diagnostizierte AHF abzuziehen, die nicht unter die Definition von kAHF im Sinne des Berichts fielen, sowie weitere 22, die in der (unsystematischen) Nachbeobachtung auffielen.

d: 27 AHF berichtet, davon 3 keine kAHF im Sinne des Berichts

kAHF: kritischer angeborener Herzfehler

Studiendesign B

Die übrigen eingeschlossenen Studien folgten dem Studiendesign B. Bei fehlender systematischer Nachbeobachtung der Testnegativen wurde auf eine Auswertung im VOPT-Design zurückgegriffen.

Zur Bewertung der Effektivität des Pulsoxymetrie-Screenings wurde in der US-amerikanischen Studie **Koppel 2003** in zwei Krankenhäusern die diagnostische Güte des Pulsoxymetrie-Screenings (Indextest I) als eine klinische Routineuntersuchung ergänzender Test untersucht (Tabelle 20). Neugeborenen im Alter von 24 bis 103 Stunden wurde der Sauerstoffsättigungswert an einem Fuß abgenommen (postductal). (Siehe auch das Flussdiagramm in Anhang F, Abbildung 4.) In den Häusern wurde die Messung zu unterschiedlichen Zeitpunkten durchgeführt, daher rührt die große Zeitspanne des Messzeitpunkts (siehe Tabelle 22). Der Trennwert für einen positiven Test lag bei ≤ 95 %. Dem Indextest I vorgeschaltet war die klinische Routineuntersuchung (Indextest II). Indextest-I-positiv getestete Neugeborene wurden echokardiografisch nachuntersucht, Referenztest der Testnegativen war die Nachbeobachtung (Tabelle 21). Die Nachbeobachtung der Indextest-I-Negativen erfolgte insbesondere anhand einer Registeranalyse. Die Autoren weisen explizit auf die Unvollständigkeit des Registers hin; nach ihrer Einschätzung ist das Register lediglich zu 87 % vollständig. Mit diesen Fehlzahlen geht eine hohe Ergebnisunsicherheit einher.

Eingeschlossen wurden ausschließlich asymptomatische Neugeborene; ausgeschlossen wurden diejenigen Neugeborenen, bei denen bereits pränatal ein kAHF diagnostiziert worden war oder die im Indextest II aufgefallen waren. In dieser Studie haben 11 281 Neugeborene beide Indextests erhalten; darunter wurden 5 mit kAHF gefunden. Der Anteil pränatal mit kAHF identifizierter Kinder in Höhe von 39 % (Tabelle 22) übersteigt leicht den deutschen Durchschnitt (siehe Kapitel 1).

Riede 2010 (Deutschland) zielte auf die Reduktion der diagnostischen Lücke ab (hierzu siehe Kapitel 1) mittels Bewertung der diagnostischen Güte von Pulsoxymetrie und klinischer Untersuchung (Tabelle 20).

Hierfür wurde bei Neugeborenen, die zuvor den Indextest II (U1 gemäß den Kinder-Richtlinien des G-BA) durchlaufen hatten, im Alter von 24 bis 72 Stunden der Sauerstoffsättigungswert an einem Fuß genommen. Der Trennwert lag bei persistierend ≤ 95 % nach 2 Messungen. Positive Untersuchungsergebnisse aus dem Indextest I wurden echokardiografisch verifiziert oder verworfen. Die Nachbeobachtung von Neugeborenen mit negativem Ergebnis des Pulsoxymetrie-Screenings beinhaltete die Erfassung aller in den an der multizentrischen Studie teilnehmenden Zentren behandelten Kinder mit später Diagnose eines kAHF sowie eine Registeranalyse (Tabelle 21; siehe auch das Flussdiagramm in Anhang G, Abbildung 5).

Das Pulsoxymetrie-Screening (Indextest I) erhielten nur reife Neugeborene (Gestationsalter ≥ 37 Wochen) mit unauffälliger U 1. Ausgeschlossen wurden diejenigen, bei denen pränatal der Verdacht auf einen kAHF bestand oder ein solcher Herzfehler diagnostiziert worden war. In Riede 2010 wurden so 18 Neugeborene mit kAHF unter 41 442 Neugeborenen identifiziert, die beide Indextests erhalten hatten. Der Anteil an pränatal mit kAHF identifizierten Kindern in Höhe von 60 % war für Deutschland überdurchschnittlich hoch (Tabelle 22).

Studienziel bei **Ruangritnamchai 2007** in Thailand war die Bestimmung von kAHF bei klinisch unauffälligen Neugeborenen durch Pulsoxymetrie-Screening; Zielkriterium war die diagnostische Güte des Pulsoxymetrie-Screenings.

Nach vorgeschalteter klinischer Routineuntersuchung im Alter von 24 bis 48 Stunden bestimmten sie mittels Pulsoxymetrie-Screening die prä- und postductale Sauerstoffsättigung des Bluts (Fuß und rechte Hand) bei einem Trennwert von < 95 % Sauerstoffsättigung. Indextest-I-Positive wurden echokardiografisch untersucht. Die Nachbeobachtung der Testnegativen erfolgte in einem Zeitraum von 1 Woche bis 2 Monaten nach negativem Indextest I (Tabelle 21). (Siehe auch das Flussdiagramm in Anhang G, Abbildung 6.)

Eingeschlossen wurden auch hier ausschließlich klinisch unauffällige Neugeborene; Ausschlusskriterium für den Indextest I war dementsprechend ein positives Untersuchungsergebnis aus Indextest II. Das Pulsoxymetrie-Screening durchliefen 1847 Neugeborene, 1 davon hatte einen kAHF. Zum Anteil pränatal mit kAHF identifizierter Kinder machen die Autoren keine Angaben (Tabelle 22).

In **Tautz 2010**, einer weiteren deutschen Studie, wurde mittels Bestimmung der diagnostischen Güte der Pulsoxymetrie die Effektivität und Praktikabilität des Pulsoxymetrie-Screenings für die sehr frühe Entdeckung von AHF untersucht (Tabelle 20).

Wie in Riede 2010 wurde bei Neugeborenen nach vorgeschaltetem Indextest II (U1 gemäß den Kinder-Richtlinien des G-BA) der Sauerstoffsättigungswert an einem Fuß genommen (Indextest I). In dieser Studie lag der Messzeitpunkt zwischen 6 und 36 Lebensstunden. Der Trennwert liegt bei Tautz 2010 bei persistierend < 95 % Sauerstoffsättigung (2 Messungen, Kontrollmessung nach 4 bis 6 Stunden). Positive Untersuchungsergebnisse aus dem Indextest I wurden echokardiografisch nachuntersucht, ebenso Neugeborene mit einer initialen Sättigung < 90 % (Tabelle 21). (Siehe auch das Flussdiagramm in Anhang G, Abbildung 7.)

Getestet wurden 3364 asymptotische Neugeborene, die keines der definierten Ausschlusskriterien aufwiesen (Gestationsalter ≤ 35 Wochen, Überweisung in die Pädiatrie, künstliche Beatmung, pränatal diagnostizierter Herzfehler). 8 Neugeborene hatten einen kAHF. Der Anteil pränatal mit kAHF identifizierter Kinder war mit 0,03 % im Vergleich zum für Deutschland angenommenen Durchschnitt sehr niedrig (Tabelle 22).

Tabelle 20: Allgemeine Studiencharakteristika (Studiendesign B)

Studie	Studienland	Studiendesign	Studienziel	Zielkriterium	Evidenzstufe ^a	Verzerrungspotenzial
Koppel 2003	USA	prospektive Studie	Effektivität des Pulsoxymetrie-Screenings	Bewertung der diagnostischen Güte von Pulsoxymetrie	I Ib	hoch / niedrig ^b
Riede 2010	Deutschland	prospektive Studie	Reduktion der diagnostischen Lücke bei kAHF	Bewertung der diagnostischen Güte von Pulsoxymetrie und klinischer Untersuchung	I Ib	niedrig
Ruangritnamchai 2007	Thailand	prospektive Studie	Bestimmung von kAHF bei klinisch unauffälligen Neugeborenen durch Pulsoxymetrie-Screening	Bewertung der diagnostischen Güte des Pulsoxymetrie-Screenings auf kAHF bei klinisch unauffälligen Neugeborenen	I Ib	hoch
Tautz 2010	Deutschland	prospektive Studie	Effektivität und Praktikabilität des Pulsoxymetrie-Screenings für die sehr frühe Entdeckung von AHF	Bewertung der diagnostischen Güte von Pulsoxymetrie	I Ib	niedrig
<p>a: entspricht der Evidenzstufe des G-BA b: Einschätzung des Verzerrungspotenzials bei VOPT-Auswertung AHF: angeborene Herzfehler; G-BA: Gemeinsamer Bundesausschuss; kAHF: kritischer angeborener Herzfehler</p>						

Tabelle 21: Index- und Referenztest (Studiendesign B)

Studie	Indextest I: POS			Indextest II: klinische Routineuntersuchung		Referenztest(s)
	Messort	Testzeitpunkt (Mittel oder Spanne)	Trennwert ^a	Testzeit (Mittel oder in Relation zu Indextest I)	Ablauf	
Koppel 2003	Fuß	> 24 Stunden bis vor Entlassung (durchschnittlich 56,9 Stunden bei vaginaler Entbindung und 103,2 Stunden bei Kaiserschnittentbindung)	▪ ≤ 95 %	vor Indextest I	(unauffällige klinische Routineuntersuchung als Einschlusskriterium)	<ul style="list-style-type: none"> ▪ Echokardiografie der Index-I-Positiven ▪ Nachbeobachtung: klinische Nachbeobachtung (Krankenhausdaten) und Registeranalyse (New York State Congenital Malformations Registry, CMR)
Riede 2010	Fuß	24–72 Stunden	▪ persistierend (2 Messungen) ≤ 95 %	vor Indextest I	(unauffällige klinische Routineuntersuchung als Einschlusskriterium)	<ul style="list-style-type: none"> ▪ Echokardiografie der Index-I-Positiven ▪ Nachbeobachtung: Bericht beteiligter Krankenhäuser und Registeranalyse (Daten der perinatalogischen und neonatologischen Qualitätssicherung im Freistaat Sachsen) während Studienlaufzeit und darüber hinaus^b
Ruangritnamchai 2007	Fuß und rechte Hand	24–48 Stunden	▪ < 95 %	24–48 Stunden	(unauffällige klinische Routineuntersuchung als Einschlusskriterium)	<ul style="list-style-type: none"> ▪ Echokardiografie der Index-I-Positiven ▪ 1 Woche bis 2 Monate-Follow-up der Testnegativen
Tautz 2010	Fuß	6–36 Stunden	<ul style="list-style-type: none"> ▪ persistierend (2 Messungen) < 95 % ▪ initial < 90 % 	(U1 unmittelbar nach der Geburt ^b)	(unauffällige klinische Routineuntersuchung als Einschlusskriterium)	<ul style="list-style-type: none"> ▪ Echokardiografie der Index-I-Positiven

a: funktionelle Sauerstoffsättigung.
b: Keine explizite Angabe im Text; die Angabe beruht auf der Antwort auf eine Autorenanfrage.

Tabelle 22: Eingeschlossene Patienten (Studiendesign B)

Studie	Eingeschlossene Neugeborene (n)				Einschlusskriterien	Ausschlusskriterien	Neugeborene mit kAHF	Anteil pränatal identifizierter Neugeborener mit kAHF	Geschlecht [n (m/w)]	Alter (in Lebensstunden zum Testzeitpunkt)
	geeignet (n)	Indextest I (n)	Indextest II (n)	Indextests I + II (n)						
Koppel 2003	11 296 (Krankenhaus A+B)	11 281	11 296	11 281	asymptomatische Neugeborene	pränatale Diagnose eines kAHF, auffällige klinische Untersuchung ^b	5	39 % (9/23)	k. A.	
	8655 (Krankenhaus A ^a)	8642	8655	8642			2			> 24
	2641 (Krankenhaus B ^a)	2639	2641	2639			3			> 56,9–103,2
Riede 2010	48 348	41 442 ^c	42 240	41 442 ^c	Neugeborene mit Gestationsalter ≥ 37 Wochen und unauffälliger klinischer Routineuntersuchung	pränataler Verdacht auf oder Diagnose eines AHF	18	60 %	k. A.	24–72
Ruangritnamchai 2007	1855	1847	1855	1847	klinisch unauffällige Neugeborene	auffällige klinische Untersuchung ^b	1	k. A.	k. A.	24–48

(Fortsetzung)

Tabelle 22: Eingeschlossene Patienten (Studiendesign B) (Fortsetzung)

Studie	Eingeschlossene Neugeborene (n)				Einschlusskriterien	Ausschlusskriterien	Neugeborene mit kAHF	Anteil pränatal identifizierter Neugeborener mit kAHF	Geschlecht [n (m/w)]	Alter (in Lebensstunden zum Testzeitpunkt)
	geeignet (n)	Indextest I (n)	Indextest II (n)	Indextests I + II (n)						
Tautz 2010	3695	3364	3364	3364	asymptomatische Neugeborene	Frühgeborene (Gestationsalter < 35 ± 0); künstlich beatmete Neugeborene, Neugeborene mit anderen Krankheiten, die in die Pädiatrie überwiesen wurden; Neugeborene mit pränatal diagnostiziertem kAHF	8	(1/3695)		6–36
<p>a: Aufgrund der unterschiedlichen Messzeitpunkte in den Krankenhäusern sind die Ergebnisse sowohl zusammen als auch getrennt voneinander abgebildet. b: Zyanose, Tachypnoe, auffällige Atemgeräusche, Nasenflügeln, Retraktion, auffällige Herzgeräusche, auffälliges Präkordium, verminderte Pulse. c: Angabe ohne Protokollverletzer (n = 3). AHF: angeborener Herzfehler; k. A.: keine Angabe; kAHF: kritischer angeborener Herzfehler</p>										

5.4.3 Einschätzung des Verzerrungspotenzials

Die Bestimmung des Verzerrungspotenzials und der Übertragbarkeit der Ergebnisse der Primärstudien auf die Fragestellung des Berichts erfolgt auf Basis von QUADAS 2 [42] (siehe Abschnitt 4.5.2). Zu den projektspezifischen Vorgaben zur Anwendung dieses Instruments siehe Anhang D.

5.4.3.1 Verzerrungspotenzial nach QUADAS 2

Im Folgenden finden sich die Ergebnisse der Bewertung des Verzerrungspotenzials. Zusammenfassend siehe Tabelle 23 und Tabelle 24.

Studiendesign A

Die Studie Zhao 2014 weist zusammenfassend ein geringes Verzerrungspotenzial auf, basierend auf einheitlich in allen 4 Domänen als niedrig eingeschätztem Verzerrungspotenzial (siehe Tabelle 23).

Studiendesign B

Die Studie Koppel 2003 weist ein hohes Verzerrungspotenzial auf. Ausschlaggebend für diese Bewertung war, dass die Daten der Indextest-I-negativ getesteten Neugeborenen in erster Linie aus einem mit ca. 87 % als unvollständig bekannten Register herrühren. Um die Daten aus dieser Studie mit denen der anderen Studien im Studiendesign B vergleichen zu können, wurden sie zusätzlich im VOPT-Design ausgewertet. Da bei der VOPT-Auswertung die Zahlen für die negativ getesteten Neugeborenen unberücksichtigt bleiben, sind diese Werte als niedrig verzerrt einzustufen.

Bei Ruangritnamchai 2007 konnte mangels Angaben – auch auf die Autorenanfrage – eine prospektive Festlegung des Trennwerts nicht bestätigt werden. Das Verzerrungspotenzial war somit insgesamt als hoch zu bewerten (siehe Tabelle 24).

Bei Riede 2010 und Tautz 2010 konnte das Verzerrungspotenzial als niedrig bewertet werden (siehe Tabelle 24).

Tabelle 23: Verzerrungspotenzial QUADAS 2 (Studiendesign A)

Studie	Patienten- selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz- standard (Domäne 3)	Patientenfluss und zeitl. Ablauf (Domäne 4)	Zusammen- fassende Einschätzung
Zhao 2014	niedrig	niedrig	niedrig	niedrig	niedrig	niedrig

Tabelle 24: Verzerrungspotenzial QUADAS 2 (Studiendesign B)

Studie	Patienten- selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz- standard (Domäne 3)	Patienten- fluss und zeitl. Ablauf (Domäne 4)	Zusam- men- fassende Ein- schätzung
Koppel 2003	niedrig	niedrig	niedrig	hoch / niedrig ^a	niedrig	hoch / niedrig ^a
Riede 2010	niedrig	niedrig	niedrig	niedrig	niedrig	niedrig
Ruangritnamchai 2007	niedrig	hoch	niedrig	niedrig	niedrig	hoch
Tautz 2010	niedrig	niedrig	niedrig	niedrig	niedrig	niedrig

a: Einschätzung des Verzerrungspotenzials bei VOPT-Auswertung
VOPT: verification of only positive testers

5.4.3.2 Bedenken bezüglich der Übertragbarkeit nach QUADAS 2

Bei allen eingeschlossenen Studien waren Bedenken bezüglich der Übertragbarkeit nach QUADAS 2 als gering einzustufen. Zu den Einzelbewertungen siehe Tabelle 25 und Tabelle 26.

Tabelle 25: Bedenken bezüglich der Übertragbarkeit QUADAS 2 (Studiendesign A)

Studie	Patienten- selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz- standard (Domäne 3)	Zusammenfassende Einschätzung
Zhao 2014	gering	gering	gering	gering	gering

Tabelle 26: Bedenken bezüglich der Übertragbarkeit QUADAS 2 (Studiendesign B)

Studie	Patienten- selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz- standard (Domäne 3)	Zusammen- fassende Einschätzung
Koppel 2003	gering	gering	gering	gering	gering
Riede 2010	gering	gering	gering	gering	gering
Ruangritnamchai 2007	gering	gering	gering	gering	gering
Tautz 2010	gering	gering	gering	gering	gering

5.5 Ergebnisse zu den Zielgrößen

Zielgrößen für die Nutzenbewertung des Pulsoxymetrie-Screenings waren die personenbezogenen Daten zur Berechnung der diagnostischen Güte in Hinblick auf die Entdeckung von KAHF (siehe Abschnitt 4.2.4).

In Tabelle 27 bis Tabelle 30 sind die Ergebnisse zu den Zielgrößen zusammengefasst.

Studiendesign A

Aus der Studie von Zhao 2014 konnten Sensitivität und Spezifität nicht berechnet werden. Die von den Autoren eingeholten Daten zu den negativen Befunden beruhen auf einer klinischen Nachuntersuchung nach 6 Wochen und Elternfeedback für den Fall eines Ereignisses. Eine vollständige Erhebung aller Ereignisse in dem geforderten Mindestnachbeobachtungszeitraum von 6 Monaten ist damit nicht gegeben. Angaben zu richtig- und falsch-negativen Ergebnissen der Indextests sind daher nicht belastbar.

Bei Zhao 2014 wurden unter 120 707 Neugeborenen 17 mit kAHF identifiziert, die in der klinischen Untersuchung nicht aufgefallen waren. Damit liegt der Anteil der im Pulsoxymetrie-Screening zusätzlich gefundenen Neugeborenen mit kAHF im Verhältnis zur gescreenten Population bei 0,014 %. Um 1 zusätzliches Neugeborenes mit kAHF zu identifizieren, mussten in dieser Studie 7100 Neugeborene pulsoxymetrisch gescreent werden (Number Needed to Screen [NNS]) (Tabelle 27).

Der PPV des Pulsoxymetrie-Screenings liegt bei 31,8 % (95 %-KI: [27,8 %; 36,0 %]). Das heißt, dass davon auszugehen ist, dass bei knapp einem Drittel aller mittels Pulsoxymetrie positiv getesteten Neugeborenen tatsächlich ein kAHF vorliegt. Viel niedriger ist dagegen der PPV der klinischen Untersuchung, der nur bei 6,1 % liegt (95 %-KI: [5,3 %; 6,9 %]). Dieser sehr niedrige Wert schlägt sich auf den PPV bei Kombination beider Indextests nieder (Pulsoxymetrie-Screening oder klinische Untersuchung positiv: believe the positive testers). Dieser ist nur wenig höher als der des Indextests II allein und liegt bei 6,2 % (95 %-KI: [5,4 %; 7,0 %]) (Tabelle 28). Bei einer dem Studiendesign B analogen Auswertung der Daten (Pulsoxymetrie-Screening nur der Neugeborenen mit unauffälligem klinischem Untersuchungsergebnis) liegt der PPV des Pulsoxymetrie-Screenings bei 8,6 % (95 %-KI: [5,1; 13,5]). Für die Berechnung dieses Werts wird die Anzahl der durch Pulsoxymetrie-Screening nach klinischer Untersuchung zusätzlich gefundenen Neugeborenen mit kAHF (n = 17) zugrunde gelegt.

Unter den 352 falsch-positiven Befunden aus dem Pulsoxymetrie-Screening finden sich 71 Herzfehler, die als „major CHD“ eingeordnet wurden, die aber nicht unter die berichtseigene Definition von kAHF fallen (VSD, PDA und AVSD). Ferner wurden weitere Herzfehler entdeckt, die nicht unter die Definition von kAHF fallen, sowie andere Erkrankungen, wie Lungenhochdruck und andere Lungenerkrankungen, Infektionen und „preterm births“ (siehe oben, Tabelle 15). Nach Autorenangaben lag der Anteil behandlungsbedürftiger Erkrankungen unter den falsch-positiven Befunden, die durch auffällige Indextests entdeckt wurden, bei 71 %.

Tabelle 27: Anzahl zusätzlich gefundener Neugeborener mit kAHF (Studiendesign A)

Studie	Anzahl gescreenter Neugeborener	Anzahl identifizierter Neugeborener mit kAHF^a	Anzahl durch Pulsoxymetrie-Screening identifizierter Neugeborener mit kAHF	Anzahl durch klinische Routineuntersuchung identifizierter Neugeborener mit kAHF	Anzahl durch Pulsoxymetrie-Screening zusätzlich identifizierter Neugeborener mit kAHF	Anteil durch Pulsoxymetrie-Screening zusätzlich identifizierter Neugeborener mit kAHF / Anzahl gescreenter Neugeborener in %	NNS (bei Add-on-Auswertung)
Zhao 2014	120 707	222	164	205	17	0,014	7100

kAHF: kritischer angeborener Herzfehler; NNS: Number Needed to Screen
a: Absolute Anzahl im Rahmen der Studie mittels klinischer Untersuchung und Pulsoxymetrie-Screening identifizierter Neugeborener mit kAHF. Diese Zahl entspricht nicht der Prävalenz, da keine adäquate Nachbeobachtung stattgefunden hat.

Tabelle 28: Ergebnisse (Studiendesign A)

Studie	n	Indextest	Trennwert	Referenztest	Trennwert	RP	FN	FP	RN	PPV ^a in % [95 %-KI]
Zhao 2014	120 707	Pulsoxymetrie-Screening	< 90 % oder persistierend 90–< 95 % oder Differenz > 3 %	Echokardiografie der Indextest-Positiven	jeder kAHF	164	- ^b	352	- ^b	31,8 [27,8; 36,0]
	120 707	klinische Untersuchung	4 Indikatoren: Familien-Anamnese, spezielle Gesichtszüge, Herzgeräusch, extrakardiale Fehlbildungen			205	- ^b	3180	- ^b	6,1 [5,3; 6,9]
	120 707	Pulsoxymetrie-Screening + klinische Untersuchung	Indextest I positiv oder Indextest II positiv			222 ^c	- ^b	3360	- ^b	6,2 [5,4; 7,0]
	120 707	Pulsoxymetrie-Screening	Indextest I positiv (bei Indextest II negativ als Einschlusskriterium) (Add-on-Auswertung)			17 ^d	- ^b	180 ^d	- ^b	8,6 [5,1; 13,5]
<p>a: eigene Berechnung von PPV mit zugehörigem Konfidenzintervall auf Basis der Vierfeldertafeln b: Referenzstandard nicht adäquat erhoben c: 284 abzüglich 62 von Indextest I und II entdeckten nicht kritischen AHF d: eigene Berechnung kAHF: kritischer angeborener Herzfehler</p>										

Studiendesign B

Koppel 2003 berichtet für das Pulsoxymetrie-Screening eine Sensitivität von 60,0 % [14,7 %; 94,7 %) und Spezifität von 100 % (95 %-KI: [100 %; 100 %]) (zusammenfassende Auswertung beider Krankenhäuser, Tabelle 30). Diese Zahlen weisen allerdings ein hohes Verzerrungspotenzial auf (siehe Abschnitt 5.4.3.1).

Ähnlich lauten die Daten für Sensitivität und Spezifität aus Riede 2010. Die Sensitivität wird hier mit 77,8 % (95 %-KI: [52,4 %; 93,6 %]) angegeben, die Spezifität liegt bei 99,9 % (95 %-KI: [99,9 %; 99,9 %]).

Der Anteil der im Pulsoxymetrie-Screening zusätzlich gefundenen Neugeborenen mit kAHF im Verhältnis zur Anzahl gescreenter Neugeborener liegt zwischen 0,027 % und 0,238 % (Tabelle 29). In Koppel 2003 wurden unter 11 281 Neugeborenen 3 mit kAHF gefunden, die in der klinischen Untersuchung nicht aufgefallen waren (0,027 %). Bei Riede 2010 waren es 14 von 41 442 (0,034 %), in Ruangritnamchai 2007 1 von 1847 (0,054 %) und in Tautz 2010 waren es 8 von 3364 (0,238 %). In den Studien von Koppel und Riede sind weitere Kinder mit kAHF berichtet, die erst im Rahmen der Nachbeobachtung identifiziert wurden. Ein Vergleich dieser Zahlen zeigt, dass in Koppel 2003 von 5 Kindern mit kAHF 3 durch das zusätzliche Pulsoxymetrie-Screening identifiziert wurden (60 %); bei Riede 2010 waren es 14 von 18 Kindern mit kAHF (77,8 %). In den übrigen Studien hat keine adäquate Nachbeobachtung stattgefunden, so dass der Anteil durch Pulsoxymetrie gefundener Neugeborener im Verhältnis zu tatsächlich mit kAHF erkrankten Kindern nicht nachvollzogen werden kann.

Um ein zusätzliches Neugeborenes mit kAHF zu finden, mussten in den Studien zwischen 421 und 3761 Neugeborene pulsoxymetrisch untersucht werden (Tabelle 29).

Die PPV der Studien im Studiendesign B weisen ebenfalls eine hohe Spannweite auf; sie liegen zwischen 25,9 % (95 %-KI: [15,0 %; 39,7 %]) in der Studie Riede 2010 und 75,0 % (95 %-KI: [19,4 %; 99,4 %]) in Koppel 2003 (Tabelle 30). Ungeachtet dieser Spannweite wurden in allen eingeschlossenen Studien durch das Pulsoxymetrie-Screening Neugeborene mit kAHF gefunden, die in der klinischen Untersuchung unauffällig geblieben waren.

Unter den falsch-positiven Befunden fanden sich auch in diesen Studien nach Angaben der Autoren behandlungsbedürftige kardiologische und nicht kardiologische Erkrankungen; deren Anteil im Verhältnis zu den falsch-positiven Befunden beträgt bei Riede 2010 70 %. Die Autoren der anderen Studien geben als Ursache für jedes falsch-positive Untersuchungsergebnis eine behandlungsbedürftige Erkrankung an (siehe oben Tabelle 16).

Die Ergebnisse im Einzelnen sind in Tabelle 30 abgebildet.

Tabelle 29: Anzahl zusätzlich gefundener Neugeborener mit kAHF (Studiendesign B)

Studie	Anzahl gescreenter Neugeborener	Anzahl identifizierter Neugeborener mit kAHF	Anzahl durch Indextest I zusätzlich identifizierter Neugeborener mit kAHF	Anteil durch Indextest I zusätzlich identifizierter Neugeborener mit kAHF / Anzahl gescreenter Neugeborener in %	NNS	Pränatale Entdeckungsrate
Koppel 2003	11 281	5 ^a	3	0,027	3761	39 %
Riede 2010	41 442	18 ^a	14	0,034	2961	60 %
Ruangritnamchai 2007	1847	1 ^b	1	0,054	1847	k. A.
Tautz 2010	3364	8 ^b	8	0,238	421	0,03 %

k. A.: keine Angabe; kAHF: kritischer angeborener Herzfehler; NNS: Number Needed to Screen

a: Absolute Anzahl im Rahmen der Studie mittels klinischer Untersuchung, Pulsoxymetrie-Screening und Nachbeobachtung identifizierter Neugeborener mit kAHF. Zahl entspricht nicht der Prävalenz, da die Anzahl der im Rahmen der klinischen Untersuchung identifizierten Kinder mit kAHF nicht eingeflossen ist.

b: Absolute Anzahl im Rahmen der Studie mittels klinischer Untersuchung und Pulsoxymetrie-Screening identifizierter Neugeborener mit kAHF. Diese Zahl entspricht nicht der Prävalenz, da die Anzahl der im Rahmen der vorgeschalteten klinischen Untersuchung identifizierten Kinder mit kAHF nicht bekannt ist und keine adäquate Nachbeobachtung stattgefunden hat.

Tabelle 30: Ergebnisse (Studiendesign B)

Studie	n	Index-test	Trennwert	Referenzstandard	Trennwert	RP	FN	FP	RN	PPV ^a in % [95 %-KI] / Sensitivität, Spezifität in % [95 %-KI]
Koppel 2003	11 281 (Krankenhäuser A + B)	Pulsoxymetrie - Screening	≤ 95 % (postductal; Messzeitpunkt > 24 Stunden)	Echokardiografie der Indextest-I-Positiven; Nachbeobachtung	jeder kAHF	3	2	1	11 275	75,0 [19,4; 99,4] / Se: 60,0 [14,7; 94,7], Sp: 100 ^c [100; 100]
	1					1	0	8640	100 [2,5; 100] / Se: 50,0 [1,3; 98,7], Sp: 100 [100; 100]	
	2					1	1	2635	66,7 [9,4; 99,2] / Se: 66,7 [9,4; 99,2], Sp: 100 [99,8; 100]	
Riede 2010	41 442 ^d	Pulsoxymetrie - Screening	≤ 95 % persistierend (2 Messungen, postductal; Messzeitpunkt > 24 Stunden)	Echokardiografie der Indextest-I-Positiven; Nachbeobachtung	jeder kAHF	14	4	40	41 384	25,9 [15,0; 39,7] / Se: 77,8 [52,4; 93,6], Sp: 99,9 [99,9; 99,9]
Ruangritnamchai 2007	1847	Pulsoxymetrie - Screening	95 % (prä- und postductal; Messzeitpunkt > 24 Stunden)	Echokardiografie der Indextest-I-Positiven	jeder kAHF	1	- ^e	2	- ^e	33,3 [0,8; 90,6]
Tautz 2010	3364	Pulsoxymetrie - Screening	< 90 % initial; < 95 % persistierend (2 Messungen, postductal; Messzeitpunkt 6–36 Stunden)	Echokardiografie der Indextest-I-Positiven	jeder kAHF	8	- ^e	10	- ^e	44,4 [21,5; 69,2]

(Fortsetzung)

Tabelle 30: Ergebnisse (Studiendesign B) (Fortsetzung)

a: eigene Berechnung von PPV, Sensitivität und Spezifität mit zugehörigen Konfidenzintervallen auf Basis der Vierfeldertafeln
b: Aufgrund der unterschiedlichen Messzeitpunkte in den Krankenhäusern sind die Ergebnisse sowohl zusammen als auch getrennt voneinander abgebildet.
c: Die Autoren geben die Spezifität mit 99,95 % an; der hier wiedergegebene Wert von 100 % ($11\,275 / (11\,275 + 2)$) beruht auf eigener Berechnung.
d: Auswertung ohne Protokollverletzer (n = 3).
e: Referenztest nicht adäquat erhoben
kAHF: kritischer angeborener Herzfehler

5.5.1 Meta-Analysen

Wegen unterschiedlicher Studienpopulationen waren Meta-Analysen nicht angezeigt.

5.5.2 Sensitivitätsanalysen

In einzelnen Studien wurden Sensitivitätsanalysen durchgeführt, die mangels Publikation einer ausreichenden, den Ergebnissen zugrunde liegenden Datenbasis im vorliegenden Bericht nicht verwertet werden konnten. So wurden beispielsweise mehrfach statistisch signifikante Unterschiede in einzelnen Subgruppen berichtet, ohne die Häufigkeiten in den einzelnen Feldern der Kontingenztafel anzugeben. Damit sind diese Ergebnisse nicht ausreichend nachvollziehbar. Da in der Nutzenbewertung besonderes Augenmerk auf die Screeningstrategie zu richten war, werden die Ergebnisse aus diesen Analysen gleichwohl berichtet.

In Bezug auf das **Alter zum Zeitpunkt der Messung** berichtete Zhao 2014 dann eine statistisch signifikante Verringerung des Anteils falsch-positiver Befunde an allen zusätzlich auffälligen Diagnosen in Abhängigkeit des Alters, wenn das Alter als stetige Variable berücksichtigt wurde (OR [95 %-KI]: 0,67 [0,57; 0,78]; $p < 0,001$). Kein statistisch signifikanter Unterschied zeigte sich dagegen, wenn dichotomisiert, also die Sauerstoffsättigung nach den ersten 24 Lebensstunden mit der Messung zwischen 6 bis 24 Lebensstunden verglichen wurde (OR [95 %-KI]: 0,90 [0,69; 1,18]; $p = 0,46$). Die Sensitivität der Pulsoxymetrie veränderte sich bei Berücksichtigung des Alters zum Zeitpunkt der Messung als stetig im Zeitverlauf nicht (OR [95 %-KI]: 0,66 [0,37; 1,20]; $p = 0,46$). Tautz 2010 wies auf den Einfluss des Messzeitpunkts auf den Anteil falsch-negativer Befunde hin. Die im ersten Zeitintervall (6 bis 12 Lebensstunden) berichteten falsch-negativen Testergebnisse wurden in späteren Kontrollmessungen (12 bis 24, 24 bis 36 Lebensstunden) korrigiert.

Von einem Einfluss der **Anzahl der Messungen** auf das Ergebnis berichteten Riede 2010 und Tautz 2010. Wenn das Screening-Protokoll bei initial positivem Ergebnis (mindestens) eine Kontrollmessung vorsieht, sei eine Senkung der falsch-positiv-Rate zu erzielen. Bei Riede 2010 sank die Anzahl falsch-positiver Befunde von 65 auf 54 bei einer Kontrollmessung nach 1 Stunde; Tautz 2010 berichtete von 85 Messungen unter dem Trennwert, von denen nach 4 bis 6 Stunden nur 7 bestätigt wurden.

5.5.3 Subgruppenmerkmale und andere Effektmodifikatoren

Subgruppenanalysen wurden in den eingeschlossenen Studien nicht durchgeführt.

5.6 Zusammenfassung der Beleglage

Die Gesamtschau der insgesamt 6 eingeschlossenen Studien ergibt einen Anhaltspunkt für einen Nutzen der Pulsoxymetrie als ein die klinische Routineuntersuchung ergänzendes Screening auf kAHF.

Die Studie de-Wahl Granelli 2009 hat ein hohes Verzerrungspotenzial und weist damit eine geringe qualitative Ergebnissicherheit auf. Ein Zusatznutzen des Pulsoxymetrie-Screenings

kann aus dieser Studie alleine nicht abgeleitet werden. Gleichwohl stützen die Ergebnisse dieser Studie die Hypothese, dass ein die klinische Routineuntersuchung ergänzendes Pulsoxymetrie-Screening die krankheits-spezifische Morbidität bei Neugeborenen mit kAHF verringert (Endpunkt schwere präoperative Azidose (OR: 0,268 [0,110; 0,654], $p = 0,003$)). Auch bei Zugrundelegung der gesamten Geburtskohorte zur Berechnung des Gruppenunterschieds ist ein positiver Einfluss des Pulsoxymetrie-Screenings auf die krankheitsspezifische Morbidität zu sehen.

Vor diesem Hintergrund sind die Ergebnisse aus den Studien zur diagnostischen Güte zu sehen, die aufgrund der angenommenen günstigeren Prognose einer früheren Erkennung von kAHF in die Nutzenbewertung einbezogen wurden. In 2 Studien zur diagnostischen Güte wurden Daten zur Sensitivität (Koppel 2003: 60,0 % (95 %-KI: [14,7 %; 94,7 %]), Riede 2010: 77,8 % (95 %-KI: [52,4 %; 93,6 %]) und Spezifität berichtet (Koppel 2003: 100 % (95 %-KI: [100 %; 100 %]), Riede 2010: 99,9 % (95 %-KI: [99,9 %; 99,9 %])). Aus allen eingeschlossenen Studien zur diagnostischen Güte konnten die NNS und der PPV abgeleitet werden. Die Studienautoren berichten übereinstimmend, dass mit dem Pulsoxymetrie-Screening Neugeborene mit kAHF entdeckt wurden, die in der klinischen Routineuntersuchung unauffällig waren.

Die Ergebnisse haben eine große Spannbreite. Die NNS reicht von 421 bis 7100 – um 1 Neugeborenes mit kAHF zu finden, mussten in den Studien jeweils zwischen 421 und 7100 Neugeborene zusätzlich zur klinischen Routineuntersuchung pulsoxymetrisch gescreent werden. Der PPV liegt zwischen 25,9 % und 75 %: Ein Viertel bis drei Viertel der im Pulsoxymetrie-Screening auffälligen Neugeborenen hatten tatsächlich einen kAHF.

Aus dem PPV folgt umgekehrt, dass ebenfalls ein Viertel bis drei Viertel der im Pulsoxymetrie-Screening auffälligen Neugeborenen keinen kAHF hatten. In diesen Fällen handelt es sich um falsch-positive Befunde; die Vornahme unnötiger Therapie ist möglich. Die Studienautoren geben an, dass falsch-positive Befunde in 70 bis 100 % der Fälle auf andere Erkrankungen zurückzuführen waren. Die Vor- und Nachteile einer Entdeckung solcher Pathologien bereits im asymptomatischen Stadium bleiben in den eingeschlossenen Studien insgesamt ungeklärt.

Zur Größenordnung und zu den Folgen falsch-negativer Untersuchungsergebnisse aus dem Pulsoxymetrie-Screening lassen die eingeschlossenen Studien keine Aussage zu.

6 Diskussion

Ziel der vorliegenden Untersuchung war die Nutzenbewertung des Pulsoxymetrie-Screenings auf kAHF bei Neugeborenen als Add-on-Test zum bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) im Vergleich zum bisherigen diagnostischen Standard ohne Pulsoxymetrie-Screening. Auch das Pulsoxymetrie-Screening als Add-on im Vergleich mit dem bisherigen diagnostischen Standard mit opportunistischer Anwendung der Pulsoxymetrie sollte untersucht werden.

Studienpool und Qualität der Daten

Der Studienpool von 6 Studien beinhaltet 1 vergleichende Interventionsstudie und 5 Studien zur diagnostischen Güte.

Die Suche nach systematischen Übersichten brachte eine Arbeit hervor, in der vergleichende Interventionsstudien ausgewertet wurden: Die Arbeit von Wennerholm und Kollegen [60] berücksichtigte neben der hier eingeschlossenen Studie von de-Wahl Granelli 2009 als 2. Studie die von Meberg und Kollegen 2009 [61]. Diese Studie war für den vorliegenden Bericht auszuschließen, weil darin nur diejenigen Neugeborenen eine klinische Untersuchung erhalten haben, die im Pulsoxymetrie-Screening unauffällig waren. Außerdem haben nur 99 von 324 positiv getesteten Neugeborenen den Referenztest Echokardiografie erhalten [61,62].

Der Pool an Studien zur diagnostischen Güte fällt im Vergleich zu anderen systematischen Übersichtsarbeiten deutlich kleiner aus. Das liegt in erster Linie daran, dass anders als in jenen Arbeiten für den vorliegenden Bericht der Nutzen des Pulsoxymetrie-Screenings als eine den derzeitigen Standard ergänzende Untersuchung zu prüfen war. Anhang E (Tabelle 31) bietet einen Überblick über die jeweils eingeschlossenen Studien nebst Angabe über Ein- oder Ausschluss in den vorliegenden Bericht.

Ergebnisse zum Zusatznutzen des Pulsoxymetrie-Screenings

Aus der Interventionsstudie de-Wahl Granelli 2009 allein konnte kein Nutzen der Pulsoxymetrie als die klinische Routineuntersuchung ergänzendes Screening für asymptotische Neugeborene abgeleitet werden. Die Studie wies eine geringe qualitative Ergebnissicherheit auf. Daneben fielen die unterschiedlichen Prävalenzen an kAHF in Interventions- und Kontrollgruppe auf, deren Ursache aus den Daten nicht hervorgeht. Trotz hierin begründeter Zweifel an der internen Validität der Studie deuten die Ergebnisse darauf hin, dass ein die klinische Routineuntersuchung ergänzendes Pulsoxymetrie-Screening auf kAHF die mit unentdeckten kAHF assoziierte Morbidität verringern könnte und damit die dem Bericht zugrundeliegende Annahme, dass eine frühere versus eine spätere Diagnostik und Intervention mit einer besseren Prognose für Kinder mit kAHF einhergehen. Dafür spricht, dass für die Interventionsgruppe statistisch signifikant geringere Morbidität bei Kindern mit kAHF berichtet wurde (schwere präoperative Azidose, $p = 0,0025$). Auch bei Zugrundelegung der gesamten Geburtskohorte zur Berechnung des Gruppenunterschieds ist ein positiver Einfluss des Pulsoxymetrie-Screenings auf die krankheitsspezifische Morbidität zu sehen.

Wegen der unterstellten und in der Studie von de-Wahl Graneli 2009 bestätigten Annahme, dass eine frühere versus spätere Diagnostik und Intervention bei Neugeborenen mit kAHF mit einer besseren Prognose einhergehen, wurden in die Nutzenbewertung auch Studien zur diagnostischen Güte einbezogen. Diese berichten für das Pulsoxymetrie-Screening eine hohe Spezifität bei hoher Sensitivität; wobei für diese Berechnung nur die Studien von Koppel 2003 und Riede 2010 herangezogen werden konnten. Das Verzerrungspotenzial für die Ergebnisse von Koppel 2003 war als hoch einzuschätzen; die Ergebnisse zu Sensitivität und Spezifität sind jedoch ähnlich wie jene von Riede 2010. Die 5 eingeschlossenen Studien zur diagnostischen Güte zeigen übereinstimmend, dass das Pulsoxymetrie-Screening als Add-on zur klinischen Untersuchung zusätzliche Kinder mit kAHF identifiziert hat. Insgesamt wurden unter 178 641 Neugeborenen 43 Neugeborene (0,024 %) mit einem kAHF gefunden, die in der klinischen Untersuchung nicht aufgefallen waren und ohne diese zusätzliche Untersuchung erst zu einem späteren Zeitpunkt mit schwerwiegender Manifestation des Herzfehlers entdeckt worden wären. Zu beachten ist, dass der Anteil entdeckter Fälle von der Vollständigkeit der Nachbeobachtung und somit der zutreffenden Identifizierung falsch-negativer Untersuchungsergebnisse abhängig ist. Das kann für die Studien von Koppel 2003 und Riede 2010 angenommen werden. Im Übrigen hat keine adäquate Nachbeobachtung stattgefunden, weshalb auf eine VOPT-Auswertung zurückgegriffen wurde.

Das bei den Studien zur diagnostischen Güte überwiegend vorgefundene niedrige Verzerrungspotenzial und das Nichtvorliegen von Bedenken bezüglich der Übertragbarkeit lassen darauf schließen, dass vergleichbare Ergebnisse im deutschen Versorgungskontext erzielt werden können.

Falsch-positive Untersuchungsergebnisse

Alle eingeschlossenen Studien berichten falsch-positive Befunde. Ihr Anteil im Verhältnis zur Anzahl gescreenter Neugeborener liegt zwischen 0,089 ‰ in Koppel 2003 und 2,972 ‰ bei Tautz 2010.

Auf diese falsch-positiven Befunde ist besonderes Augenmerk zu richten, weil in solchen Fällen grundsätzlich die Möglichkeit unnötiger Therapie besteht (siehe Kapitel 4). Hier ist zu unterscheiden, ob das Pulsoxymetrie-Screening ein gesundes Neugeborenes findet, oder aber ob der Grund für das Untersuchungsergebnis eine andere als die Zielerkrankung kAHF ist. Für den ersten Fall gilt, dass durch die echokardiografische Nachuntersuchung aller pulsoxymetrisch positiv getesteten Neugeborenen sichergestellt ist, dass falsch-positive Befunde herausgefiltert werden. Bis zu diesem Zeitpunkt ist gleichwohl eine in diesen Fällen unnötige medikamentöse Behandlung wahrscheinlich. Im zweiten Fall dagegen sind die Folgen unklar. Die Autoren der eingeschlossenen Studien bewerten solche falsch-positiven Befunde, die auf andere Erkrankungen als kAHF zurückzuführen sind, durchweg positiv. Ihr Anteil bezogen auf alle falsch-positiven Befunde ist bei Riede 2010 mit 70 % angegeben, bei Zhao 2014 mit 71 %. In den übrigen eingeschlossenen Studien wurden sogar alle falsch-positiven Untersuchungsergebnisse auf behandlungsbedürftige Erkrankungen zurückgeführt. Zu beachten ist, dass aufgrund des Designs in diesen Studien sämtliche berichteten falsch-

positiven Befunde aus dem Pulsoxymetrie-Screening hervorgegangen sind. Darunter finden sich andere, nicht kritische Herzfehler, Lungenentzündungen, Lungenhochdruck, Neugeborenensepsis sowie weitere Pathologien.

Ob es tatsächlich von Vorteil ist, Neugeborene mit den hier genannten anderen Pathologien als kAHF bereits asymptomatisch zu identifizieren, oder ob den Neugeborenen hieraus Nachteile erwachsen, lässt sich anhand der gesichteten Studien nicht beurteilen. Ein Schaden des Pulsoxymetrie-Screenings könnte in einer Stigmatisierung durch falsch-positive Befunde und gegebenenfalls unnötiger abklärender Diagnostik oder Therapie bestehen, der auch die Eltern betreffen kann, indem diese durch das Ergebnis beunruhigt werden. Ewer und Kollegen haben diese Frage mittels Fragebogenerhebung untersucht [18]. Sie haben 30 Tage nach der Geburt (Mittelwert; Interquartilsabstand 12 bis 58 Tage; Zeitpunkt der Beantwortung des Fragebogens) keine signifikante Erhöhung der Besorgnis (anxiety) bei Eltern von Neugeborenen mit falsch-positivem Pulsoxymetrie-Ergebnis im Vergleich zu Eltern richtig-negativ getesteter Neugeborener vorgefunden, wohl aber einen statistisch signifikant höheren Depression-Score und eine geringere Gesamtzufriedenheit mit dem Pulsoxymetrie-Screening.

Falsch-negative Untersuchungsergebnisse

Das Pulsoxymetrie-Screening erzeugt auch falsch-negative Testergebnisse. Gleichwohl sind Neugeborene mit falsch-negativem Ergebnis aus dem Pulsoxymetrie-Screening nicht benachteiligt gegenüber Neugeborenen, die kein zusätzliches Pulsoxymetrie-Screening erhalten haben. In jedem Fall erfolgt die klinische Routineuntersuchung, die ihrerseits bei positivem Befund zur weiterführenden kardiologischen Untersuchung führt.

Vergleich der Screeningstrategien

Angesichts voneinander abweichender Anwendungsweisen des Pulsoxymetrie-Screenings (siehe Anhang G) konnten die Ergebnisse der Studien nicht gepoolt werden; und auch für eine Aussage zum zu bevorzugenden Ablauf des Screenings bieten die Studien keine hinreichende Datenbasis. Dies betrifft mögliche Effektmodifikatoren wie den genauen Zeitpunkt der Messung nach Geburt, die Anzahl und den zeitlichen Abstand weiterer Messungen, die Anzahl der Messpunkte (prä- und postductal) und den Trennwert für die Sauerstoffsättigung. Gleichwohl bleibt festzuhalten, dass alle Studien, unabhängig von den Unterschieden in der Anwendung des Pulsoxymetrie-Screenings, einen konsistenten Effekt zeigten. Unter anderem zu Ermittlung der optimalen Anwendung des Pulsoxymetrie-Screenings hat das UK National Screening Committee die Durchführung eines Pilot-Projekts in 6 Krankenhäusern beschlossen [63].

Ergebnisse und Schlussfolgerungen anderer systematischer Übersichten

Die Ergebnisse aus den 6 im Rahmen der Recherche identifizierten systematischen Übersichten (siehe Anhang C) stützen die des vorliegenden Berichts. Die folgende Darstellung fokussiert exemplarisch auf die beiden umfassendsten Übersichtsarbeiten von Wennerholm 2011 [60] und Thangaratinam 2012 [27].

In Wennerholm 2011 wurden 2 Interventionsstudien in Hinblick auf patientenrelevante Endpunkte ausgewertet (Meberg 2009 [61] und de-Wahl Granelli 2009 [17]).

Für den Endpunkt Mortalität konnte aus Meberg 2009 mangels Ereignis weder in der Studien- noch in der Kontrollgruppe das relative Risiko (RR) geschätzt werden. Aus der Studie de-Wahl Granelli 2009 wollten die Autoren aufgrund der wenigen Ereignisse keine Schlüsse ziehen. Die Evidenz nach GRADE bewerteten Wennerholm und Kollegen für den Endpunkt Mortalität mit sehr gering (very low), für den Endpunkt Morbidität mit gering (low). Sie kamen ebenso wie der vorliegende Bericht zu dem Ergebnis, dass die zusätzliche Anwendung des Pulsoxymetrie-Screenings das Risiko schwerer Azidosen reduziert [60].

In Hinblick auf die diagnostische Güte des Pulsoxymetrie-Screenings kommen Wennerholm und Kollegen zu dem Ergebnis, dass sowohl das Pulsoxymetrie-Screening isoliert als auch die Kombination aus Pulsoxymetrie-Screening und klinischer Untersuchung eine gute diagnostische Güte aufweisen. Der Evidenzgrad wird als moderat (moderate) beschrieben.

Aus 2 Studien schließen die Autoren, dass die kombinierte Diagnostik besser sei als das Pulsoxymetrie-Screening oder die klinische Untersuchung allein [17,62]. Statistische Analysen wurden nicht durchgeführt.

Anders dagegen in der aktuellen Übersichtsarbeit von Thangaratnam und Kollegen aus 2012 [27]. Die Arbeit schließt 13 Studien mit insgesamt 229 421 Neugeborenen ein. Im Studienpool sind 2 auch im vorliegenden Bericht eingeschlossene Studien [20,54], daneben unter anderem nicht vergleichende Studien sowie Abstract-Publikationen und Fall-Kontroll-Studien.

Thangaratnam berichtet eine Sensitivität des Pulsoxymetrie-Screenings auf kritische Herzfehler bei Neugeborenen als Replacement (und nicht wie im vorliegenden Bericht als Add-on zur klinischen Routineuntersuchung) von 76,5 % (95 %-KI: [67,7 %; 83,5 %]) und Spezifität von 99,9 % (95 %-KI: [99,7 %; 99,9 %]). Ähnliche Werte wurden auch in den eingeschlossenen Studien des vorliegenden Berichts berichtet.

Der Meta-Analyse liegt eine Definition von kAHF zugrunde, die der des vorliegenden Berichts sehr nahe kommt (Definition von [18,19]). Wie im vorliegenden Bericht mühten sich die Autoren um den Abgleich der Daten zu gefundenen Herzfehlern in Hinblick auf ihre Definition. Anders als im vorliegenden Bericht hingegen gab es bezüglich des Referenztests keine konkreten Vorgaben, sodass dieser in den eingeschlossenen Studien höchst unterschiedlich ausfiel. Bei den Ergebnissen dieser Meta-Analyse ist daher zu beachten, dass aufgrund der unterschiedlichen Referenztests die Anzahl der falsch-negativ getesteten Neugeborenen möglicherweise zu niedrig geschätzt ist. Die Autoren verweisen mit Blick auf ihre Funnel-Plot-Analyse auf die Möglichkeit, dass die Meta-Analyse größere Effekte zeigt als in Wirklichkeit gegeben.

Thangaratinam und Kollegen bescheinigen der Pulsoxymetrie eine „überzeugende Evidenz“ für die Implementierung als Screening in die klinische Praxis.

Ergebnisse und Schlussfolgerungen von Leitlinien und Empfehlungen

Das Screening auf kAHF mittels Pulsoxymetrie ist nach einigen nationalen und internationalen Leitlinien und Empfehlungen ohne Leitliniencharakter bereits Bestandteil der Neugeborenenversorgung. Neben 2 deutschen Leitlinien [14,15] beinhalten weitere europäische und internationale Leitlinien und Empfehlungen das Pulsoxymetrie-Screening als zusätzliche Untersuchung auf kAHF im Rahmen der Neugeborenenversorgung [14-16,59,64-67].

Das Spektrum der Anwendungsweise des Pulsoxymetrie-Screenings dieser Leitlinien und Empfehlungen ist ebenso weit wie das der in diesen Bericht eingeschlossenen Studien. Eine allgemeine, von Experten und Fachgesellschaften bevorzugte Screeningstrategie lässt sich auch hieraus nicht ableiten. Im Gegenteil wird weiterer Forschungsbedarf zur Bestimmung der optimalen Strategie gesehen [59]. Einen Überblick über die gesichteten Leitlinien und Empfehlungen nebst Screeningstrategie bietet Tabelle 33 in Anhang E.

Die Verfasser der Leitlinien und der Empfehlungen bescheinigen dem Pulsoxymetrie-Screening als Add-on ausreichende Evidenz für eine Implementierung in die routinemäßige Neugeborenenversorgung. Überwiegend wird dies damit begründet, dass die Pulsoxymetrie schon lange, zum Beispiel im NICU-Bereich, etabliert sei und dass es sich um eine einfache und kostengünstige Maßnahme handele. Bemerkenswert ist, dass lediglich in 2 Empfehlungen ohne Leitliniencharakter ein möglicher Schaden durch falsch-positive und falsch-negative Befunde erwähnt wird [59,63]. Genannt werden die Beunruhigung der Eltern und die Kostenverursachung durch zusätzliche Echokardiografien sowie der Trugschluss, dass ein negatives Pulsoxymetrie-Screening einen kAHF ausschließe. Eine weitere Empfehlung weist auf die Grenzen des Pulsoxymetrie-Screenings hin [67]. Diese 3 Empfehlungen raten zu entsprechender Elterninformation. Fast alle Leitlinien und Empfehlungen beschreiben falsch-positive kAHF-Befunde, die auf anderen Pathologien hinweisen, als Vorteil.

Um eine standardisierte Implementierung des Pulsoxymetrie-Screenings auf kAHF im europäischen Raum bemüht sich eine Gruppe internationaler Neonatologen und Kardiologen um de-Wahl Granelli und Ewer. Auf der Grundlage „hinreichend robuster Evidenz“ drängen sie auf die Formulierung entsprechender politischer Vorgaben [13].

Gesamtschau und Einordnung der Ergebnisse des Berichts

Die Bestimmung des Sauerstoffgehalts des Blutes mittels Pulsoxymetrie ist ein seit Langem etabliertes Verfahren; neu ist allerdings die Anwendung bei Neugeborenen als Screening auf kAHF.

Zusammenfassend kann davon ausgegangen werden, dass ein die klinische Routineuntersuchung ergänzendes Pulsoxymetrie-Screening zur früheren Entdeckung von Neugeborenen mit kAHF führt. Gesunde Neugeborene erfahren durch dieses zusätzliche

Screening zumindest keinen unmittelbaren Nachteil, weil es eine nicht invasive, das Neugeborene nicht belastende und einfach durchzuführende Untersuchung ist.

Unklar sind die Folgen falsch-positiver Testergebnisse, die allein durch das die klinische Routineuntersuchung ergänzende Pulsoxymetrie-Screening erzeugt wurden und die auf andere Erkrankungen und Ereignisse als kAHF zurückzuführen sind. Offen ist, inwieweit es sich zum Beispiel auswirkt, wenn bei einem Neugeborenen aufgrund pulsoxymetrischer Untersuchung ein nicht kritischer Herzfehler gefunden wird, der mit keinerlei klinischen Auswirkungen einhergeht. In solchen Fällen wird unnötige abklärende Diagnostik folgen. Weiter könnte es sein, dass allein das Wissen um einen Herzfehler zu einer Belastung zunächst der Eltern, später auch des Kindes führt. Denkbar ist auch, dass es zu einer unbegründeten Schonung des Kindes kommt, wenn es etwa um die Teilnahme an Sport- oder Freizeitangeboten geht [68]. Gravierende Nachteile sind denkbar, wenn es aufgrund von falsch-positivem Befund zu einer möglicherweise zu diesem frühen Zeitpunkt unnötigen Therapie kommt, die ihrerseits beispielsweise als invasiver Eingriff oder medikamentöse Behandlung Belastungen oder unerwünschten Nebenwirkungen auslösen.

Schädliche Folgen falsch-negativer Testergebnisse sind dann denkbar, wenn Eltern oder medizinisches Personal bei anderweitigen Anzeichen für einen kAHF diesem Verdacht wegen des unauffälligen Testergebnisses nicht mit weiterer abklärender Diagnostik nachgehen. In diesem Zusammenhang ist zu beachten, dass insbesondere isolierte Linksherzobstruktionen (CoA, IAA) mittels Pulsoxymetrie nur schwer zu identifizieren sind [17,53,54,62]. Bei der Wahl Granelli 2009 wird vor diesem Hintergrund angeregt, den Perfusionsindex (peripheral perfusion index, PPI) als weitere Diagnostik zu verwenden; einen Wert, den Pulsoxymeter der neuen Generation bereits liefern. Bei falsch-negativem pulsoxymetrischen Untersuchungsergebnis wäre es mithin möglich, dass notwendige medizinische Interventionen verzögert werden. Da das Pulsoxymetrie-Screening die U1 / U2 ergänzen, nicht aber ersetzen soll, sollte von dieser Konstellation in der Praxis nicht auszugehen sein. Ein negatives Testergebnis des Pulsoxymetrie-Screenings schließt die Neugeborenen nicht von den übrigen Untersuchungen entsprechend dem bisherigen diagnostischen Standard aus. Insofern ist durch falsch-negative Ergebnisse des Pulsoxymetrie-Screenings kein gravierender Schaden zu erwarten.

Um möglichen Nachteilen falsch-positiver und falsch-negativer Testergebnisse vorzubeugen, sollte das Angebot des Pulsoxymetrie-Screenings mit einer aussagekräftigen und verständlichen schriftlichen Aufklärung der Eltern und auch des medizinischen Personals über die Grenzen dieses Screenings einhergehen. Hierzu gehört insbesondere die Information, dass das Pulsoxymetrie-Screening auch dazu geeignet ist, nicht kritische AHF sowie weitere Neugeborenen-Erkrankungen zu entdecken, die gegebenenfalls bereits asymptomatisch behandlungsbedürftig sein können, aber nicht müssen. Außerdem muss darüber informiert werden, dass das Pulsoxymetrie-Screening keine abschließende Sicherheit über das Nichtvorliegen eines kAHF bieten kann. So ist bekannt, dass bestimmte kAHF durch das Screening nur schwierig zu identifizieren sind.

Insbesondere zur Beantwortung der Frage nach der optimalen Screeningstrategie sowie den Langzeitfolgen ist weitere Forschung wünschenswert. Auf Basis der diesem Bericht zugrundeliegenden Daten ist davon auszugehen, dass der Zusatznutzen einen möglichen Schaden überwiegt und damit die in diesem Zusammenhang angestellten Überlegungen für eine Einführung des Pulsoxymetrie-Screenings sprechen, eingebettet in die erforderlichen Informationen über die Grenzen dieses Tests.

7 Fazit

Das Pulsoxymetrie-Screening als ein den bisherigen diagnostischen Standard (U1- und U2-Screening oder vergleichbare klinische Untersuchungen) ergänzendes Screening auf kAHF bei asymptomatischen Neugeborenen zeigt einen Anhaltspunkt für einen Nutzen. Dieses Ergebnis ist auf 1 vergleichende Interventionsstudie und 5 Studien zur diagnostischen Güte zurückzuführen: Die 5 Studien zur diagnostischen Güte zeigen übereinstimmend, dass mit dem Pulsoxymetrie-Screening zusätzliche Neugeborene mit kAHF entdeckt wurden, die in der klinischen Routineuntersuchung unauffällig waren. Das Ergebnis der vergleichenden Interventionsstudie unterstützt die Hypothese, dass die krankheitsspezifische Morbidität bei Neugeborenen mit kAHF durch das zusätzliche Pulsoxymetrie-Screening verringert werden kann. Der Endpunkt schwere präoperative Azidose bei Neugeborenen mit kAHF zeigte einen statistisch signifikanten Unterschied zugunsten des Pulsoxymetrie-Screenings. Unsicherheit besteht hinsichtlich einer Nutzen-Schaden-Abwägung für die nicht intendierten Befunde, die mit der Pulsoxymetrie zusätzlich erkannt werden.

8 Liste der eingeschlossenen Studien

De-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganäs L et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39 821 newborns. *BMJ* 2009; 338: a3037.

Koppel RI, Druschel CM, Carter T, Goldberg BE, Mehta PN, Talwar R et al. Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003; 111(3): 451-455.

Riede FT, Worner C, Dahnert I, Mockel A, Kostelka M, Schneider P. Effectiveness of neonatal pulse oximetry screening for detection of critical congenital heart disease in daily clinical routine: results from a prospective multicenter study. *Eur J Pediatr* 2010; 169(8): 975-981.

Ruangritnamchai C, Bunjapamai W, Pongpanich B. Pulse oximetry screening for clinically unrecognized critical congenital heart disease in the newborns. *Images Paediatr Cardiol* 2007; 9(1): 10-15.

Tautz J, Merkel C, Loersch F, Egen O, Hägele F, Thon HM et al. Implication of pulse oxymetry screening for detection of congenital heart defects. *Klin Padiatr* 2010; 222(5): 291-295.

Zhao QM, Ma XJ, Ge XL, Liu F, Yan WL, Wu L et al. Pulse oximetry with clinical assessment to screen for congenital heart disease in neonates in China: a prospective study. *Lancet* 23.04.2014 [Epub ahead of print].

9 Literatur

1. Lindinger A, Schwedler G, Hense HW. Prevalence of congenital heart defects in newborns in Germany: results of the first registration year of the PAN Study (July 2006 to June 2007). *Klin Padiatr* 2010; 222(5): 321-326.
2. Deutsches Institut für Medizinische Dokumentation und Information. Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision; German modification; Version 2013 [online]. 21.09.2012 [Zugriff: 03.09.2013]. URL: <http://www.dimdi.de/static/de/klassi/icd-10-gm/kodesuche/onlinefassungen/htmlgm2013/index.htm>.
3. Lentze MJ, Schaub J, Schulte FJ, Spranger J (Ed). *Pädiatrie: Grundlagen und Praxis*. Berlin: Springer 2003.
4. Riede FT, Hamsch J, Dähnert I, Schneider P, Janousek J. Kritische angeborene Herzfehler: pathophysiologische Grundlagen und praktische Hinweise zur Akutdiagnostik und -therapie bei Neugeborenen. *Kinder- und Jugendmedizin* 2009; 9(3): 133-139.
5. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the American Heart Association and American Academy of Pediatrics. *Circulation* 2009; 120(5): 447-458.
6. Gemeinsamer Bundesausschuss. Richtlinien des Bundesausschusses der Ärzte und Krankenkassen über die ärztliche Betreuung während der Schwangerschaft und nach der Entbindung („Mutterschafts-Richtlinien“) [online]. 18.07.2013 [Zugriff: 04.10.2013]. URL: http://www.g-ba.de/downloads/62-492-773/RL_Mutter-2013-07-18_2013-09-20.pdf.
7. Friedberg MK, Silverman NH, Moon-Grady AJ, Tong E, Nourse J, Sorenson B et al. Prenatal detection of congenital heart disease. *J Pediatr* 2009; 155(1): 26-31.e1.
8. Sklansky MS, Berman DP, Pruetz JD, Chang RK. Prenatal screening for major congenital heart disease: superiority of outflow tracts over the 4-chamber view. *J Ultrasound Med* 2009; 28(7): 889-899.
9. Gemeinsamer Bundesausschuss. Richtlinien des Bundesausschusses der Ärzte und Krankenkassen über die Früherkennung von Krankheiten bei Kindern bis zur Vollendung des 6. Lebensjahres („Kinder-Richtlinien“) [online]. 16.12.2010 [Zugriff: 05.09.2013]. URL: http://www.g-ba.de/downloads/62-492-506/RL_Kinder_2010-12-16.pdf.
10. Salyer JW. Neonatal and pediatric pulse oximetry. *Respir Care* 2003; 48(4): 386-396.
11. Thangaratinam S, Daniels J, Ewer AK, Zamora J, Khan KS. Accuracy of pulse oximetry in screening for congenital heart disease in asymptomatic newborns: a systematic review. *Arch Dis Child Fetal Neonatal Ed* 2007; 92(3): F176-F180.
12. Valmari P. Should pulse oximetry be used to screen for congenital heart disease? *Arch Dis Child Fetal Neonatal Ed* 2007; 92(3): F219-F224.

13. Ewer AK, De-Wahl Granelli A, Manzoni P, Sanchez Luna M, Martin GR. Pulse oximetry screening for congenital heart defects. *Lancet* 2013; 382(9895): 856-857.
14. Herting E, Vetter K, Gonser M, Bassler D, Hentschel R, Groneck P. Betreuung von gesunden reifen Neugeborenen in der Geburtsklinik [online]. 13.09.2012 [Zugriff: 05.09.2013]. URL: http://www.awmf.org/uploads/tx_szleitlinien/024-0051_S2k_Betreuung_von_gesunden_reifen_Neugeborenen_2012-10.pdf.
15. Franke J, Haas NA. S 2 Leitlinie Pädiatrische Kardiologie: Abklärung einer Zyanose [online]. 01.10.2011 [Zugriff: 03.09.2013]. URL: http://www.kinderkardiologie.org/Leitlinien/16%20LL%20Abklaerung%20einer%20Zyanose%20DGPK%20lay%20out161211%20_2_.pdf.
16. Arlettaz R, Bauersfeld U. Empfehlungen zum neonatalen Screening kongenitaler Herzfehler [online]. 2006 [Zugriff: 06.09.2013]. URL: http://www.neonet.ch/assets/pdf/Empfehlungen_POx_CHD_SGN_d_edited.pdf.
17. De-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganäs L et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39 821 newborns. *BMJ* 2009; 338: a3037.
18. Ewer AK, Furnston AT, Middleton LJ, Deeks JJ, Daniels JP, Pattison HM et al. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness. *Health Technol Assess* 2012; 16(2): v-xiii, 1-184.
19. Ewer AK, Middleton LJ, Furnston AT, Bhojar A, Daniels JP, Thangaratnam S et al. Pulse oximetry screening for congenital heart defects in newborn infants (PulseOx): a test accuracy study. *Lancet* 2011; 378(9793): 785-794.
20. Riede FT, Worner C, Dahnert I, Mockel A, Kostelka M, Schneider P. Effectiveness of neonatal pulse oximetry screening for detection of critical congenital heart disease in daily clinical routine: results from a prospective multicenter study. *Eur J Pediatr* 2010; 169(8): 975-981.
21. Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. *Arch Dis Child Fetal Neonatal Ed* 2008; 93(1): F33-F35.
22. Arlettaz R, Bauschatz AS, Monkhoff M, Essers B, Bauersfeld U. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *Eur J Pediatr* 2006; 165(2): 94-98.
23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.1. Köln: IQWiG; 2013. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf.
24. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.

25. Mellander M. Diagnosis and management of life-threatening cardiac malformations in the newborn. *Semin Fetal Neonatal Med* 2013; 18(5): 302-310.
26. Dhandayuthapani G, Chakrabarti S, Ranasinghe A, Hunt L, Grant D, Martin RP et al. Short-term outcome of infants presenting to pediatric intensive care unit with new cardiac diagnoses. *Congenit Heart Dis* 2010; 5(5): 444-449.
27. Thangaratinam S, Brown K, Zamora J, Khan KS, Ewer AK. Pulse oximetry screening for critical congenital heart defects in asymptomatic newborn babies: a systematic review and meta-analysis. *Lancet* 2012; 379(9835): 2459-2464.
28. Schultz AH, Localio AR, Clark BJ, Ravishankar C, Videon N, Kimmel SE. Epidemiologic features of the presentation of critical congenital heart disease: implications for screening. *Pediatrics* 2008; 121(4): 751-757.
29. Oster ME, Lee KA, Honein MA, Riehle-Colarusso T, Shin M, Correa A. Temporal trends in survival among infants with critical congenital heart defects. *Pediatrics* 2013; 131(5): e1502-e1508.
30. Brown JW, Park HJ, Turrentine MW. Arterial switch operation: factors impacting survival in the current era. *Ann Thorac Surg* 2001; 71(6): 1978-1984.
31. Hayes DA, Jones S, Quaegebeur JM, Richmond ME, Andrews HF, Glickstein JS et al. Primary arterial switch operation as a strategy for total correction of taussig-bing anomaly: a 21-year experience. *Circulation* 2013; 128(26 Suppl 1): S194-S198.
32. Knowles R, Griebisch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2005; 9(44): 1-152, iii-iv.
33. Hoffman JI. It is time for routine neonatal screening by pulse oximetry. *Neonatology* 2011; 99(1): 1-9.
34. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. *Heart* 2006; 92(9): 1298-1302.
35. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3; current step 4 version [online]. 30.11.1995 [Zugriff: 09.03.2012]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
36. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.

37. 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.
38. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
39. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Ann Intern Med* 2013; 158(7): 544-554.
40. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.
41. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003; 138(1): 40-44.
42. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-536.
43. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
44. Lange S. The all randomized/full analysis set (ICH E9): may patients be excluded from the analysis? *Drug Inf J* 2001; 35(3): 881-891.
45. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
46. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.
47. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the Bernoulli parameter. *Am Stat* 1996; 50(1): 63-68.
48. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990.
49. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006; 59(12): 1331-1332.
50. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-64.
51. Hotelling H. The generalization of student's ratio. *Annals of Mathematical Statistics* 1931; 2(3): 360-378.

52. Centers for Disease Control and Prevention. Screening for critical congenital heart defect [online]. 13.05.2013 [Zugriff: 21.06.2013]. URL: <http://www.cdc.gov/ncbddd/pediatricgenetics/CCHDscreening.html>.
53. Zhao QM, Ma XJ, Ge XL, Liu F, Yan WL, Wu L et al. Pulse oximetry with clinical assessment to screen for congenital heart disease in neonates in China: a prospective study. *Lancet* 23.04.2014 [Epub ahead of print].
54. Koppel RI, Druschel CM, Carter T, Goldberg BE, Mehta PN, Talwar R et al. Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003; 111(3): 451-455.
55. Ruangritnamchai C, Bunjapamai W, Pongpanich B. Pulse oximetry screening for clinically unrecognized critical congenital heart disease in the newborns. *Images Paediatr Cardiol* 2007; 9(1): 10-15.
56. Tautz J, Merkel C, Loersch F, Egen O, Hägele F, Thon HM et al. Implication of pulse oxymetry screening for detection of congenital heart defects. *Klin Padiatr* 2010; 222(5): 291-295.
57. De-Wahl Granelli A, Mellander M, Sunnegardh J, Sandberg K, Ostman-Smith I. Screening for duct-dependant congenital heart disease with pulse oximetry: a critical evaluation of strategies to maximize sensitivity. *Acta Paediatr* 2005; 94(11): 1590-1596.
58. Deutsche Gesellschaft für Neonatologie und Pädiatrische Intensivmedizin, Deutsche Gesellschaft für Gynäkologie und Geburtshilfe, Deutsche Gesellschaft für Kinderheilkunde und Jugendmedizin, Deutsche Gesellschaft für Perinatale Medizin. Verlegung von Früh- und Reifgeborenen in Krankenhäuser der adäquaten Versorgungsstufe [online]. 04.2013 [Zugriff: 22.09.2014]. URL: http://www.awmf.org/uploads/tx_szleitlinien/024-0021_S1_Verlegung_von_Früh-und_Reifgeborenen_in_Krankenhäuser_der_adäquaten_Versorgungsstufe_2013-04.pdf.
59. Kemper AR, Mahle WT, Martin GR, Cooley WC, Kumar P, Morrow WR et al. Strategies for implementing screening for critical congenital heart disease. *Pediatrics* 2011; 128(5): e1259-e1267.
60. Wennerholm UB, Fassoulas A, Daxberg EL, Hafström O, Liljegren A, Samuelsson O et al. Pulse oximetry (POX) screening for congenital heart defects in newborns [online]. 06.2011 [Zugriff: 23.07.2014]. (HTA-Rapporter; Band 36). URL: <http://www.sahlgrenska.se/upload/SU/HTA-centrum/HTA-rapporter/HTA-report%20POX%20%202011-06-30%20till%20publicering%202.pdf>.
61. Meberg A, Andreassen A, Brunvand L, Markestad T, Moster D, Nietsch L et al. Pulse oximetry screening as a complementary strategy to detect critical congenital heart defects. *Acta Paediatrica* 2009; 98(4): 682-686.

62. Meberg A, Brugmann-Pieper S, Due R Jr, Eskedal L, Fagerli I, Farstad T et al. First day of life pulse oximetry screening to detect congenital heart defects. *Journal of Pediatrics* 2008; 152(6): 761-765.
63. Health Service Executive, Royal College of Physicians of Ireland. Pulse oximetry testing for newborn congenital heart disease [online]. [Zugriff: 23.07.2014]. URL: <http://www.hse.ie/eng/about/Who/clinical/natclinprog/paediatricsandneonatologyprogramme/chd.pdf>.
64. Helsedirektoratet. Nytt liv og trygg barseltid for familien: nasjonal faglig retningslinje for barselomsorgen [online]. 09.04.2014 [Zugriff: 24.07.2014]. URL: http://helsedirektoratet.no/publikasjoner/nasjonal-retningslinje-nytt-liv-og-trygg-barseltid-for-familien-retningslinje-for-barselomsorgen/Publikasjoner/IS_2057_Barsel_retningslinjer_fullversjon.pdf.
65. Ministra Zdrowia. Rozporządzenie Ministra Zdrowia z dnia 23 września 2010 r.: w sprawie standardów postępowania oraz procedur medycznych przy udzielaniu świadczeń zdrowotnych z zakresu opieki okołoporodowej sprawowanej nad kobietą w okresie fizjologicznej ciąży, fizjologicznego porodu, położu oraz opieki nad noworodkiem. *Dziennik Ustaw* 2010; (187): 1259.
66. Mahle WT, Martin GR, Beekman RH 3rd, Morrow WR. Endorsement of Health and Human Services recommendation for pulse oximetry screening for critical congenital heart disease. *Pediatrics* 2012; 129(1): 190-192.
67. Fritz M, Fritsch P, Foramitti M, Simma B. Pulsoxymetriescreening bei Neugeborenen auf kritische angeborene Herzfehler: Empfehlungen der Arbeitsgruppen Pädiatrische Kardiologie und Neonatologie/ Pädiatrische Intensivmedizin der Österreichischen Gesellschaft für Kinder- und Jugendheilkunde. *Monatsschr Kinderheilkd* 2014; 162(7): 638–643.
68. Bergman AB, Stamm SJ. The morbidity of cardiac nondisease in schoolchildren. *N Engl J Med* 1967; 276(18): 1008-1013.

Anhang A – Suchstrategien**1. EMBASE****Suchoberfläche: Ovid**

- EMBASE 1974 to 2013 October 14

#	Searches
1	Infant/
2	Newborn/
3	Newborn Screening/
4	(newborn* or babies* or baby* or neonate* or neonatal* or infant* or postnatal* or birth*).ti,ab.
5	or/1-4
6	exp Congenital Heart Disease/
7	Cardiovascular Malformation/cn
8	(heart* adj3 (disease* or defect* or syndrome*)).ti,ab.
9	((cardiovascular* or cardiac* or congenital*) adj3 (anomalies* or anomaly* or malformation* or defect*)).ti,ab.
10	or/6-9
11	exp Oximetry/
12	(pulse* adj1 (oximetry* or oximetric* or oximeter*)).ti,ab.
13	((oximetric* or oximetry*) adj6 screening*).ti,ab.
14	or/11-13
15	and/5,10,14

2. MEDLINE**Suchoberfläche: Ovid**

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations October 01, 2013
- Ovid MEDLINE(R)1946 to October Week 1 2013
- Ovid MEDLINE(R) Daily Update October 01, 2013

#	Searches
1	exp Infant, Newborn/
2	Neonatal Screening/
3	(newborn* or babies* or baby* or neonate* or neonatal* or infant* or postnatal* or birth*).ti,ab.
4	or/1-3

#	Searches
5	exp Heart Defects, Congenital/
6	(heart* adj3 (disease* or defect* or syndrome*)).ti,ab.
7	((cardiovascular* or cardiac* or congenital*) adj3 (anomalies* or anomaly* or malformation* or defect*)).ti,ab.
8	or/5-7
9	exp Oximetry/
10	(pulse* adj1 (oximetry* or oximetric* or oximeter*)).ti,ab.
11	((oximetric* or oximetry*) adj6 screening*).ti,ab.
12	or/9-11
13	and/4,8,12

3. PubMed

Suchoberfläche: NLM

- PubMed – as supplied by publisher
- PubMed – in process
- PubMed – OLDMEDLINE
- PubMed – pubmednotmedline

Search	Most Recent Queries
#1	Search newborn*[tiab] OR babies*[tiab] OR baby*[tiab] OR neonate*[tiab] OR neonatal*[tiab] OR infant*[tiab] OR postnatal*[tiab] OR birth*[tiab]
#2	Search heart*[tiab] AND (disease*[tiab] OR defect*[tiab] OR syndrome*[tiab])
#3	Search (cardiovascular*[tiab] OR cardiac*[tiab] OR congenital*[tiab]) AND (anomalies*[tiab] OR anomaly*[tiab] OR malformation*[tiab] OR defect*[tiab])
#4	Search #2 OR #3
#5	Search pulse*[tiab] AND (oximetry*[tiab] OR oximetric*[tiab] OR oximeter*[tiab])
#6	Search (oximetric*[tiab] OR oximetry*[tiab]) AND screening*[tiab]
#7	Search #5 OR #6
#8	Search #1 AND #4 AND #7
#9	Search #8 NOT medline[sb]

4. The Cochrane Library**Suchoberfläche: Wiley**

Cochrane Database of Systematic Reviews (Cochrane Reviews), Issue 10, 2013

Cochrane Central Register of Controlled Trials (Clinical Trials), Issue 9, 2013

#	Searches
1	MeSH descriptor: [Infant, Newborn] explode all trees
2	MeSH descriptor: [Neonatal Screening] this term only
3	(newborn* or babies* or baby* or neonate* or neonatal* or infant* or postnatal* or birth*):ti,ab
4	#1 or #2 or #3
5	MeSH descriptor: [Heart Defects, Congenital] explode all trees
6	(heart* near/3 (disease* or defect* or syndrome*))
7	((cardiovascular* or cardiac* or congenital*) near/3 (anomalies* or anomaly* or malformation* or defect*))
8	#5 or #6 or #7
9	MeSH descriptor: [Oximetry] explode all trees
10	(pulse* near/1 (oximetry* or oximetric* or oximeter*))
11	((oximetric* or oximetry*) near/6 screening*)
12	#9 or #10 or #11
13	#4 and #8 and #12

Database of Abstracts of Reviews of Effects (Other Reviews), Issue 3, 2013

Health Technology Assessment Database (Technology Assessments), Issue 3, 2013

#	Searches
1	MeSH descriptor: [Infant, Newborn] explode all trees
2	MeSH descriptor: [Neonatal Screening] this term only
3	(newborn* or babies* or baby* or neonate* or neonatal* or infant* or postnatal* or birth*)
4	#1 or #2 or #3
5	MeSH descriptor: [Heart Defects, Congenital] explode all trees
6	(heart* near/3 (disease* or defect* or syndrome*))
7	((cardiovascular* or cardiac* or congenital*) near/3 (anomalies* or anomaly* or malformation* or defect*))
8	#5 or #6 or #7

#	Searches
9	MeSH descriptor: [Oximetry] explode all trees
10	(pulse* near/1 (oximetry* or oximetric* or oximeter*))
11	((oximetric* or oximetry*) near/6 screening*)
12	#9 or #10 or #11
13	#4 and #8 and #12

Anhang B – Liste der ausgeschlossenen Dokumente mit Ausschlussgründen**E2b**

1. Andrews JP, Ross AS, Salazar MA, Tracy NA, Burke BL Jr. Smooth implementation of critical congenital heart defect screening in a newborn nursery. *Clin Pediatr (Phila)* 13.09.2013 [Epub ahead of print].
2. Arlettaz R, Bauschatz AS, Monkhoff M, Essers B, Bauersfeld U. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *Eur J Pediatr* 2006; 165(2): 94-98.
3. Bakr AF, Habib HS. Combining pulse oximetry and clinical examination in screening for congenital heart disease. *Pediatr Cardiol* 2005; 26(6): 832-835.
4. Green E, Rosenkvist CJ. Pulse oximetry screening of newborns detects congenital heart defects: experiences from Kalmar [Schwedisch]. *Lakartidningen* 2012; 109(9-10): 481-482.
5. Meberg A, Andreassen A, Brunvand L, Markestad T, Moster D, Nietsch L et al. Pulse oximetry screening as a complementary strategy to detect critical congenital heart defects. *Acta Paediatr* 2009; 98(4): 682-686.
6. Meberg A, Brugmann-Pieper S, Due R Jr, Eskedal L, Fagerli I, Farstad T et al. First day of life pulse oximetry screening to detect congenital heart defects. *J Pediatr* 2008; 152(6): 761-765.
7. Reich JD, Connolly B, Bradley G, Littman S, Koepfel W, Lewycky P et al. Reliability of a single pulse oximetry reading as a screening test for congenital heart disease in otherwise asymptomatic newborn infants: the importance of human factors. *Pediatr Cardiol* 2008; 29(2): 371-376.
8. Reich JD, Miller S, Brogdon B, Casatelli J, Gompf TC, Huhta JC et al. The use of pulse oximetry to detect congenital heart disease. *J Pediatr* 2003; 142(3): 268-272.
9. Richmond S, Reay G, Abu Harb M. Routine pulse oximetry in the asymptomatic newborn. *Arch Dis Child Fetal Neonatal Ed* 2002; 87(2): F83-F88.
10. Rosati E, Chitano G, Dipaola L, De Felice C, Latini G. Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. *J Perinat Med* 2005; 33(5): 455-457.
11. Samuel TY, Bromiker R, Mimouni FB, Picard E, Lahav S, Mandel D et al. Newborn oxygen saturation at mild altitude versus sea level: implications for neonatal screening for critical congenital heart disease. *Acta Paediatr* 2013; 102(4): 379-384.
12. Vaidyanathan B, Sathish G, Mohanan ST, Sundaram KR, Warriar KKR, Kumar RK. Clinical screening for congenital heart disease at birth: a prospective study in a community hospital in Kerala. *Indian Pediatr* 2011; 48(1): 25-30.

E3b

1. Han LM, Klewer SE, Blank KM, Seckeler MD, Barber BJ. Feasibility of pulse oximetry screening for critical congenital heart disease at 2643-foot elevation. *Pediatr Cardiol* 2013; 34(8): 1803-1807.
2. Hoke TR, Donohue PK, Bawa PK, Mitchell RD, Pathak A, Rowe PC et al. Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatr Cardiol* 2002; 23(4): 403-409.
3. Sendelbach DM, Jackson GL, Lai SS, Fixler DE, Stehel EK, Engle WD. Pulse oximetry screening at 4 hours of age to detect critical congenital heart defects. *Pediatrics* 2008; 122(4): e815-e820.
4. Turska Kmiec A, Borszewska Kornacka MK, Błaz W, Kawalec W, Zuk M. Early screening for critical congenital heart defects in asymptomatic newborns in Mazovia province: experience of the POLKARD pulse oximetry programme 2006-2008 in Poland. *Kardiologia Pol* 2012; 70(4): 370-376.
5. Walsh W. Evaluation of pulse oximetry screening in Middle Tennessee: cases for consideration before universal screening. *J Perinatol* 2011; 31(2): 125-129.

E4b

1. Bradshaw EA, Cuzzi S, Kiernan SC, Nagel N, Becker JA, Martin GR. Feasibility of implementing pulse oximetry screening for congenital heart disease in a community hospital. *J Perinatol* 2012; 32(9): 710-715.
2. Fetter WPF, Lafeber HN, Van Lingen RA, Sauer PJJ. Pulse oxymetry in neonates [Niederländisch]. *Ned Tijdschr Geneesk* 1988; 132(18): 815-819.
3. Górska-Kot A, Błaz W, Pszeniczna E, Rusin J, Materna-Kirylyuk A, Homa E et al. Trends in diagnosis and prevalence of critical congenital heart defects in the Podkarpacie province in 2002-2004, based on data from the Polish Registry of Congenital Malformations. *J Appl Genet* 2006; 47(2): 191-194.
4. Griebisch I, Knowles RL, Brown J, Bull C, Wren C, Dezateux CA. Comparing the clinical and economic effects of clinical examination, pulse oximetry, and echocardiography in newborn screening for congenital heart defects: a probabilistic cost-effectiveness model and value of information analysis. *Int J Technol Assess Health Care* 2007; 23(2): 192-204.
5. Mai CT, Riehle-Colarusso T, O'Halloran A, Cragan JD, Onley RS, Lin A et al. Selected birth defects data from population-based birth defects surveillance programs in the United States, 2005-2009: featuring critical congenital heart defects targeted for pulse oximetry screening. *Birth Defects Res A Clin Mol Teratol* 2012; 94(12): 970-983.
6. Reich JD, Connolly B, Bradley G, Littman S, Koeppel W, Lewycky P et al. The reliability of a single pulse oximetry reading as a screening test for congenital heart disease in otherwise asymptomatic newborn infants. *Pediatr Cardiol* 2008; 29(5): 885-889.

E5b

1. Barrington KJ. Pulse oximetry before discharge from the nursery can increase detection of serious congenital heart disease. *J Pediatr* 2009; 155(1): 147-148.
2. Barrington KJ. Pulse oximetry is a cost-effective addition in screening newborns for congenital heart defects. *J Pediatr* 2012; 161(3): 569-570.
3. Centers for Disease Control and Prevention. Rapid implementation of pulse oximetry newborn screening to detect critical congenital heart defects: New Jersey, 2011. *MMWR Morb Mortal Wkly Rep* 2013; 62(15): 292-294.
4. Ewer AK, Furnston AT, Middleton LJ, Deeks JJ, Daniels JP, Pattison HM et al. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness. *Health Technol Assess* 2012; 16(2): v-xiii, 1-184.
5. Ewer AK, Middleton LJ, Furnston AT, Bhojar A, Daniels JP, Thangaratinam S et al. Pulse oximetry screening for congenital heart defects in newborn infants (PulseOx): a test accuracy study. *Lancet* 2011; 378(9793): 785-794.
6. Garg LF, Van Naarden Braun K, Knapp MM, Anderson TM, Koppel RI, Hirsch D et al. Results from the New Jersey statewide critical congenital heart defects screening program. *Pediatrics* 2013; 132(2): e314-e323.
7. Hines AJ. A nurse-driven algorithm to screen for congenital heart defects in asymptomatic newborns. *Adv Neonatal Care* 2012; 12(3): 151-157.
8. Hokanson JS. Pulse oximetry is beneficial in screening newborns for critical congenital heart disease. *J Pediatr* 2012; 160(3): 529.
9. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the AHA and AAP. *Pediatrics* 2009; 124(2): 823-836.
10. Martin GR, Bradshaw EA. Sensitivity of pulse oximetry for detection of critical congenital heart defects in newborn infants higher than that of antenatal ultrasound with few false positives. *Evid Based Med* 2012; 17(2): 57-58.
11. Mellander M. Diagnosis and management of life-threatening cardiac malformations in the newborn. *Semin Fetal Neonatal Med* 2013; 18(5): 302-310.

E7b

1. Riede FT, Dahnert I, Woerner C, Mockel A, Lorenz N, Kabus M et al. Pulsoxymetriescreening kann die diagnostische Lücke bei kritischen angeborenen Herzfehlern verkleinern: erste Ergebnisse einer prospektiven multizentrischen Feldstudie in Sachsen. *Monatsschr Kinderheilkd* 2009; 157(9): 896-902.

Anhang C – Liste der gesichteten systematischen Übersichten

1. Knowles R, Griebisch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2005; 9(44): 1-152, iii-iv.
2. Ma XJ, Huang GY. Neonatal pulse oximetry screening improves detecting of critical congenital heart disease. *Chin Med J (Engl)* 2013; 126(14): 2736-2740.
3. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the American Heart Association and American Academy of Pediatrics. *Circulation* 2009; 120(5): 447-458.
4. Thangaratinam S, Brown K, Zamora J, Khan KS, Ewer AK. Pulse oximetry screening for critical congenital heart defects in asymptomatic newborn babies: a systematic review and meta-analysis. *Lancet* 2012; 379(9835): 2459-2464.
5. Thangaratinam S, Daniels J, Ewer AK, Zamora J, Khan KS. Accuracy of pulse oximetry in screening for congenital heart disease in asymptomatic newborns: a systematic review. *Arch Dis Child Fetal Neonatal Ed* 2007; 92(3): F176-F180.
6. Wennerholm UB, Fassoulas A, Daxberg EL, Hafström O, Liljegren A, Samuelsson O et al. Pulse oximetry (POX) screening for congenital heart defects in newborns [online]. 06.2011 [Zugriff: 23.07.2014]. (HTA-Rapporter; Band 36). URL: <http://www.sahlgrenska.se/upload/SU/HTA-centrum/HTA-rapporter/HTA-report%20POX%20%202011-06-30%20till%20publicering%20.pdf>.

Anhang D – Projektspezifische Vorgaben zur Anwendung von QUADAS 2

QUADAS 2 gliedert sich in die 4 Domänen Patientenselektion, Indextest(s), Referenzstandard sowie Patientenfluss und zeitlicher Ablauf. Als Bewertungsstufen kommen für das Verzerrungspotenzial niedrig, hoch und unklar in Betracht, für Bedenken bezüglich der Übertragbarkeit gering, stark und unklar. Die Beantwortung von Indikatorfragen zu jeder Studie erlaubt die studienübergreifend möglichst konsistente Bewertung.

Innerhalb der Domänen wurden vorab projektspezifische Vorgaben formuliert, um dem Ziel der konsistenten Bewertung so umfassend wie möglich gerecht zu werden.

Domäne 1 befasst sich mit der Patientenselektion. Hier wird geprüft, ob die Auswahl der Patienten zu systematischen Verzerrungen geführt haben könnte. In Hinblick auf die Übertragbarkeit der Ergebnisse ist zu beurteilen, ob die in die Studie eingeschlossenen Patienten der Fragestellung des Berichts entsprechen.

Hier galt es die pränatale Entdeckungsrate von kAHF zusammen mit dem Ein- oder Ausschluss dieser Neugeborenen in die Studie zu beachten. Eine hohe pränatale Entdeckungsrate in Verbindung mit dem Studienausschluss vermindert die Anzahl derjenigen Neugeborenen mit kAHF, die durch Pulsoxymetrie erkannt werden können, während eine niedrige pränatale Entdeckungsrate oder der Einschluss dieser Neugeborenen diese Anzahl erhöht.

Für Deutschland wurde nach derzeitigem Standard von einer durchschnittlichen Entdeckungsrate von etwa einem Drittel aller kAHF ausgegangen (siehe Kapitel 1). Vor diesem Hintergrund könnte man Studien mit stark abweichender pränataler Entdeckungsrate „starke“ Bedenken bezüglich der Übertragbarkeit attestieren. Allerdings weist die pränatale Entdeckungsrate von kAHF innerhalb Deutschlands, also der für die Fragestellung des Berichts relevanten Population, starke Schwankungen auf. Daher sollte auf eine solche Wertung verzichtet werden. Die pränatale Entdeckungsrate von kAHF wird daher wertungsfrei berichtet.

Domäne 2 bezieht sich auf die Indextests. Für die Bewertung des Verzerrungspotenzials des Indextests I, dem Pulsoxymetrie-Screening, ist von erheblicher Bedeutung, ob der in der Studie verwendete Trennwert prospektiv festgelegt wurde. Eine nachträgliche Bestimmung erlaubt die Optimierung in Hinblick auf die zutreffende Identifizierung von kAHF und kann so zur Überschätzung der diagnostischen Güte führen. Als erforderlich für die Annahme der prospektiven Festlegung des Trennwerts wurde eine eindeutige diesbezügliche Angabe angesehen. Die Korrespondenzautoren aller eingeschlossenen Studien, in denen diese Angabe fehlt, wurden diesbezüglich angefragt. Bei fehlender anderslautender Rückmeldung blieb es bei der Feststellung der nicht prospektiven Festlegung des Trennwerts mit der Folge, dass das Verzerrungspotenzial des Indextests I als „hoch“ einzustufen war.

Bei der Frage nach dem Verzerrungspotenzial des Indextests II (klinische Routineuntersuchung) stellt sich die Problematik nicht. Für die klinische Untersuchung gibt es in der Regel

keinen vergleichbar definierten Trennwert zur Abgrenzung eines auffälligen von einem unauffälligen Untersuchungsergebnis in Hinblick auf einen KAHF. Daher sollte diese Indikatorfrage offen bleiben können. Diesbezügliche Angaben werden gleichwohl berichtet.

Das Verhältnis zwischen den Indextests bildet QUADAS 2 nicht ab. Im Verhältnis zwischen dem Pulsoxymetrie-Screening und der klinischen Routineuntersuchung hatte sich jedoch die Frage gestellt, inwiefern es sich auswirkt, wenn in einer Studie der Indextest I ohne Wissen über das Ergebnis des Indextests II (oder umgekehrt) durchgeführt wird oder aber in Kenntnis des Ergebnisses.

Jedenfalls immer dann, wenn die Studie dem Studiendesign B folgt, wird das Pulsoxymetrie-Screening in Kenntnis des Ergebnisses der klinischen Routineuntersuchung durchgeführt: Ein negatives Untersuchungsergebnis aus Indextests II ist Voraussetzung für den Studieneinschluss des Neugeborenen und damit die Anwendung des Pulsoxymetrie-Screenings. Da die Fragestellung des Berichts auf eine zusätzliche, die klinische Routineuntersuchung ergänzende Anwendung des Pulsoxymetrie-Screenings abzielt, entspricht ein solches Studiendesign der klinischen Praxis und es wurde festgelegt, dass sich die Kenntnis des klinischen Untersuchungsergebnisses aus Indextest II bei Durchführung des Indextests I nicht negativ auswirkt.

Bei der Bewertung des Verzerrungspotenzials des Referenztests (Domäne 3) wird zunächst die Wahrscheinlichkeit beurteilt, ob der Referenztest die Zielerkrankung korrekt klassifiziert. Hier war zu beachten, dass je nach Ergebnis der Indextests entweder eine Echokardiografie (bei positivem Testergebnis) oder Nachbeobachtung (bei negativem Testergebnis) zu erwarten war. Beides – Echokardiografie und Nachbeobachtung – wurde als ein Referenztest bewertet. Beide Methoden wurden als gleichermaßen geeignet angesehen, KAHF korrekt zu identifizieren; bei der Nachbeobachtung unabhängig von ihrer Dauer oder systematischen Durchführung. Diesbezügliche Bedenken finden an anderer Stelle (Domäne 4, siehe unten) Berücksichtigung.

Ferner fragt QUADAS 2, ob der Referenztest ohne Wissen der Ergebnisse der Indextests durchgeführt wurde. Da davon auszugehen war, dass die Ergebnisse der Indextests in der Regel ausschlaggebend für die Wahl des Referenztests sein würden (ausgenommen ein Design, das für jedes Neugeborene eine echokardiografische Untersuchung vorsieht), war zu erwarten, dass diese Indikatorfrage regelmäßig zu verneinen sein würde. Bei entsprechendem Studiendesign sollte sich diese Antwort daher nicht negativ auf die Bewertung des Verzerrungspotenzials auswirken.

Unter Domäne 4 sind schließlich der Patientenfluss und der zeitliche Ablauf zu bewerten.

Die erste Indikatorfrage bezieht sich auf die Angemessenheit der Zeitspanne zwischen Indextests und Referenztest. Bei Fehlen entgegenstehender Angaben in der Studienpublikation sollte regelmäßig davon ausgegangen werden, dass die Zeitspanne angemessen ist.

Ferner ist zu prüfen, ob alle Patienten einen Referenzstandard erhalten haben. Da in die Nutzenbewertung auch Studien im VOPT-Design eingeschlossen werden sollten, sollten Studien, die auf eine Verifizierung der Testnegativen verzichtet haben, als solche kenntlich gemacht werden, ohne dass sich dieses Design negativ auf die Bewertung des Verzerrungspotenzials auswirkt. Die Auswirkungen des partial verification-bias wurden damit bewusst als hinnehmbar gewertet.

Auch bei der 3. Indikatorfrage zu Domäne 4 galt es eine Besonderheit der zu bewertenden Studien zu beachten: Gefragt wird danach, ob alle Patienten denselben Referenzstandard erhalten haben. Da vorliegend 2 Methoden als ein Referenztest akzeptiert wurden, sollte sich deren Anwendung ebenfalls nicht nachteilig auf die Bewertung des Verzerrungspotenzials auswirken. Auch die Auswirkungen des differential verification-bias wurden somit bewusst als hinnehmbar gewertet.

Anhang E – Relevante Primärstudien und ihr Einschluss in systematische Übersichten (nach ausgewerteten Studien, tabellarische Übersicht)

Tabelle 31: Relevante Primärstudien und ihr Einschluss in systematische Übersichten

Übersichts- arbeiten Studien (nach Publikationsdatum)	MA 2013	Thanga- ratinam 2012	Wenner- holm 2011	Mahle 2009	Thanga- ratinam 2007	Knowles 2005	Einschluss (E) / Ausschluss (A) mit nicht erfülltem Einschluss- kriterium
Zhao 2014 ^a							E
Ewer 2011	x	x					A E4b
Riede 2010	x	x	x				E
Tautz 2010			x				E
de Wahl-Granelli 2009	x	x	x				E / A E 5b
Meberg 2009	x		x	x			A E2b
Meberg 2008		x	x				A E2b
Reich 2008	x						A E4b
Sendelbach 2008		x	x	x			A E3b
Ruangritnamchai 2007 ^a							E
Arlettaz 2006	x	x	x	x	x		A E2b
Kawalec 2006		x		x			A E7b
Bakr 2005	x	x		x	x		A E 2b
de Wahl-Granelli 2005					x		A min. EK verletzt
Rosati 2005	x	x	x	x	x		A E2b
Koppel 2003	x	x	x	x	x	x	E
Reich 2003		x		x	x	x	A E2b
Hoke 2002	x	x		x	x	x	A E3b
Richmond 2002	x	x		x	x	x	A E2b

a: in keiner der systematischen Übersichten eingeschlossen.

Anhang F– Leitlinien und Empfehlungen (tabellarische Übersicht)

Tabelle 32: Screeningempfehlungen ausgewählter Leitlinien

Leitlinie	Pulsoxymetrie			Klinische Routineuntersuchung	
	Messort	Testzeitpunkt	Trennwert	Untersuchung	Testzeitpunkt
national					
DGGG, DGKJ, DGPM, Deutscher Hebammenverband, GNPI Standard 2012, Betreuung von gesunden reifen Neugeborenen in der Geburtsklinik S2k	Fuß	2. Lebens- tag	< 95 %	U1-Erstunter- suchung	nach der Geburt
	alternativ: rechte Hand und Fuß		alternativ: Differenz > 3 %	U2-Vorsorge- untersuchung	3.–10. Lebenstag
DGPK 2011, Leitlinie pädiatrische Kardiologie S2	alle 4 Ex- tremitäten	1. Lebens- tag	<ul style="list-style-type: none"> ▪ < 90 % ▪ persistierend < 95 % (90–94 % ohne Verdacht auf Vitium: 2 Messungen mit 12-Stunden-Abstand) ▪ > 95 % 	klinische Untersuchung (Herzgeräusch)	k. A.
international					
Helsedirektoratet Norway 2014, Nytt liv og trygg barseltid for familien Nasjonal faglig retningslinje for barselomsorgen	Fuß	≥ 24 Lebens- stunden	<ul style="list-style-type: none"> ▪ persistierend < 95 % (2 Messungen im Abstand von 2–4 Stunden) 	klinische Routine- untersuchung	1.–3. Lebenstag
	alternativ: rechte Hand und Fuß				
Rozporządzenie Ministra Zdrowia ^a 2010, Standard porodu fizjologicznego.	rechter Fuß	2–24 Lebens- stunden	≤ 95 %	klinische Routine- untersuchung	nach der Geburt
				regelmäßige Bewertung des Allgemein- zustands	während der ersten 48. Lebensstunden
Swiss Society of Neonatology 2005/2006 Guidelines Empfehlungen zum neonatalen Screening kongenitaler Herzfehler	Fuß	1. Lebens- tag	<ul style="list-style-type: none"> ▪ < 90 % ▪ persistierend < 95 % (90–94 % bei 2 Messungen mit einigen Stunden Abstand) ▪ ≥ 95 % 	klinische Routine- untersuchung	2. oder 3. Lebenstag
a: Gesundheitsminister, Polen. k. A.: keine Angabe					

Tabelle 33: Internationale Screeningempfehlungen ohne Leitliniencharakter

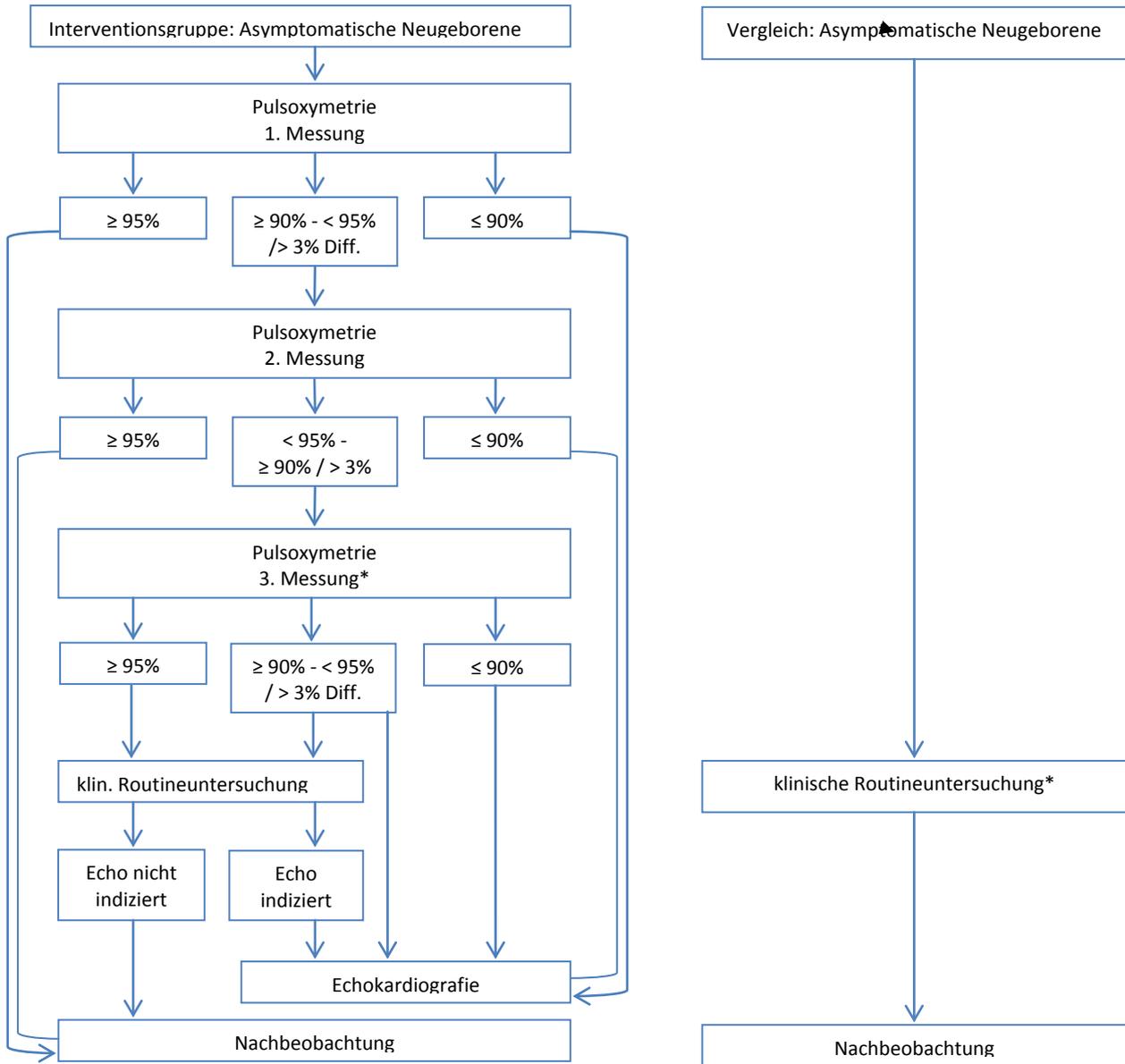
Empfehlungen	Pulsoxymetrie			Klinische Routineuntersuchung	
	Messort	Testzeitpunkt	Trennwert	Untersuchung	Testzeitpunkt
American Academy of Pediatrics (AAP) 2012, Endorsement of Health and Human Services. Recommendation for Pulse Oxymetry Screening for critical Congenital Heart Disease (unter Bezugnahme auf Kemper und Kollegen 2011)	rechte Hand und ein Fuß	> 24 Lebensstunden / bei früherer Entlassung so spät wie möglich	< 90 %, persistierend < 95 %	klinische Untersuchung	k. A.
	rechte Hand und ein Fuß		persistierend Differenz > 3 %		
Kemper et al. 2011, Strategies for Implementing Screening for Critical Congenital Heart Disease	rechte Hand und ein Fuß	2. Lebenstag / bei früherer Entlassung so spät wie möglich	<ul style="list-style-type: none"> ▪ < 90 % ▪ persistierend 90–<95 % (3 Messungen mit jeweils 1 Stunde Abstand) 	klinische Untersuchung	k. A.
	rechte Hand. und ein Fuß		<ul style="list-style-type: none"> ▪ persistierend > 3 % Differenz (3 Messungen mit jeweils 1 Stunde Abstand) 		
Health Service Executive, Royal College of Physicians of Ireland 2012, Pulse Oximetry Testing for Newborn Congenital Heart Disease	Fuß	> 24 Lebensstunden , bei früherer Entlassung unmittelbar vorher	<ul style="list-style-type: none"> ▪ < 90 % ▪ persistierend < 95 % (2 Messungen im Abstand von 1–2 Stunden) 	klinische Routineuntersuchung	k. A.
Arbeitsgruppen Pädiatrische Kardiologie und Neonatologie/ Pädiatrische Intensivmedizin der Österreichischen Gesellschaft für Kinder- und Jugendheilkunde (Österreich) 2014, Pulsoxymetriescreening bei Neugeborenen auf kritische angeborene Herzfehler	Fuß	24.–48. Lebensstunde	<ul style="list-style-type: none"> ▪ < 90 % ▪ persistierend 90–94 % (3 Messungen) im Abstand von jeweils 1 Stunde) 	klinische Observation	-
k. A.: keine Angabe					

Anhang G – Screeningstrategien der eingeschlossenen Studien

Tabelle 34: Übersicht Pulsoxymetrie-Screeningstrategien

Studie	Studien- design	pränatale Entdeckungs- rate	Mess- zeitpunkt in Stunden	Messort	Anzahl Messungen	Trennwert
Koppel 2003	B	39 %	> 24	postductal	1	≤ 95 %
Riede 2010		60 %	> 24	postductal	2 (wenn initial ≤ 95 %)	≤ 95 %
Ruangritnamchai 2007		k. A.	> 24	prä- und postductal	1	< 95 %
Tautz 2010		< 1 %	6–36 (MW: k. A.)	postductal	2 (wenn initial 90– < 95 %); 1 (wenn initial ≤ 90 %)	90–< 95 %; ≤ 90 %
Zhao 2014	A	8,7 %	6–72 (MW: 43)	prä- und postductal	2 (wenn initial 90– 95 % oder > 3 % Differenz zw. prä- und postductal); 1 (wenn initial < 90 %)	90–95 %; > 3 % Differenz; < 90 %
k. A.: keine Angabe						

Vergleichende Interventionsstudie
de-Wahl Granelli 2009



* k.A. bzgl. Referenztest bei Verdacht auf kAHF

Abbildung 3: Screeningstrategie de-Wahl Granelli 2009

Studien zur diagnostischen Güte

Koppel 2003

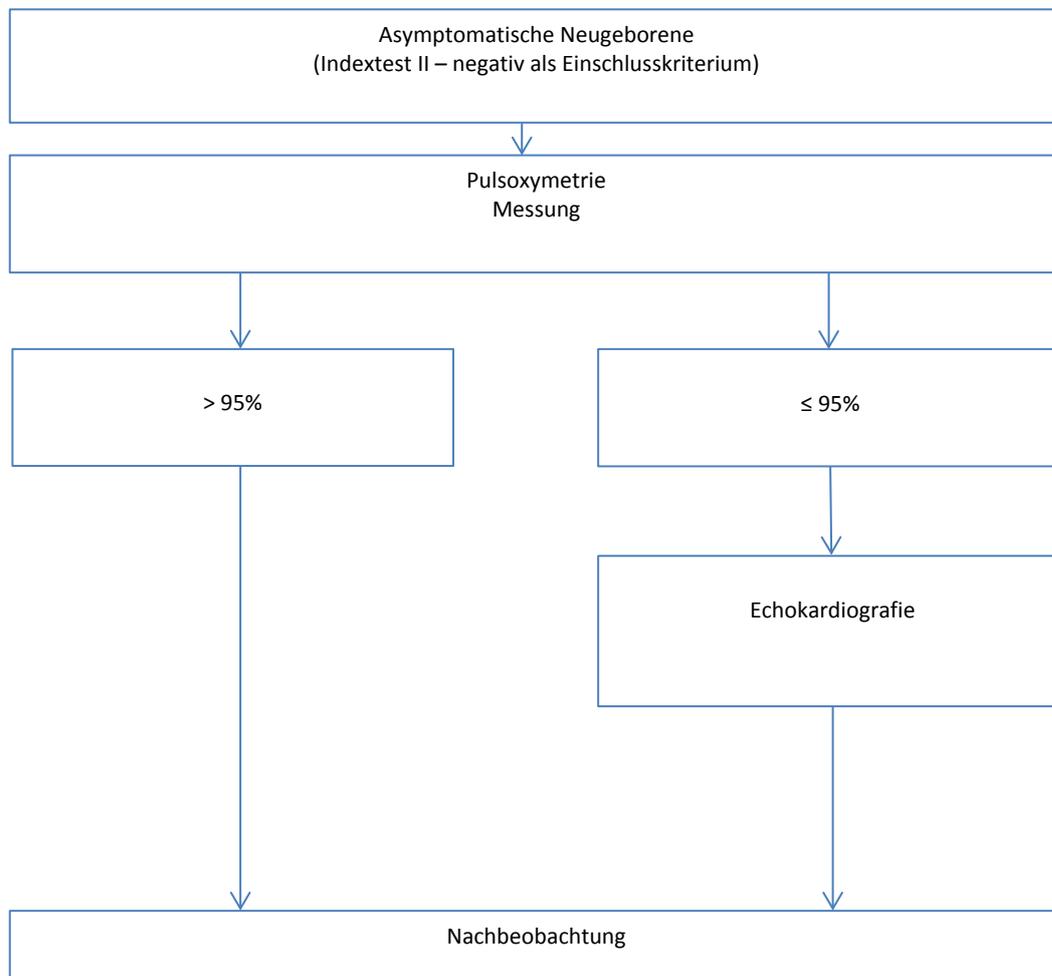


Abbildung 4: Screeningstrategie Koppel 2003

Riede 2010

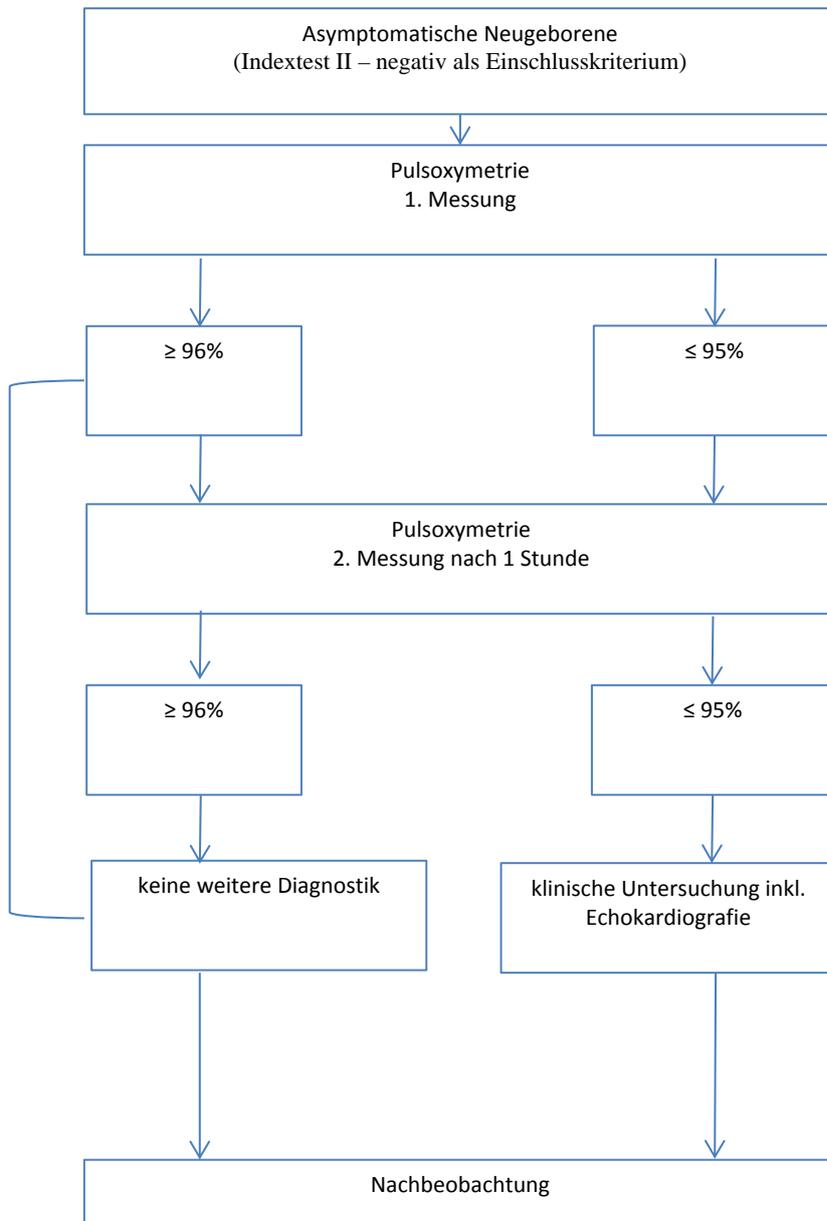


Abbildung 5: Screeningstrategie Riede 2010

Ruangritnamchai 2010

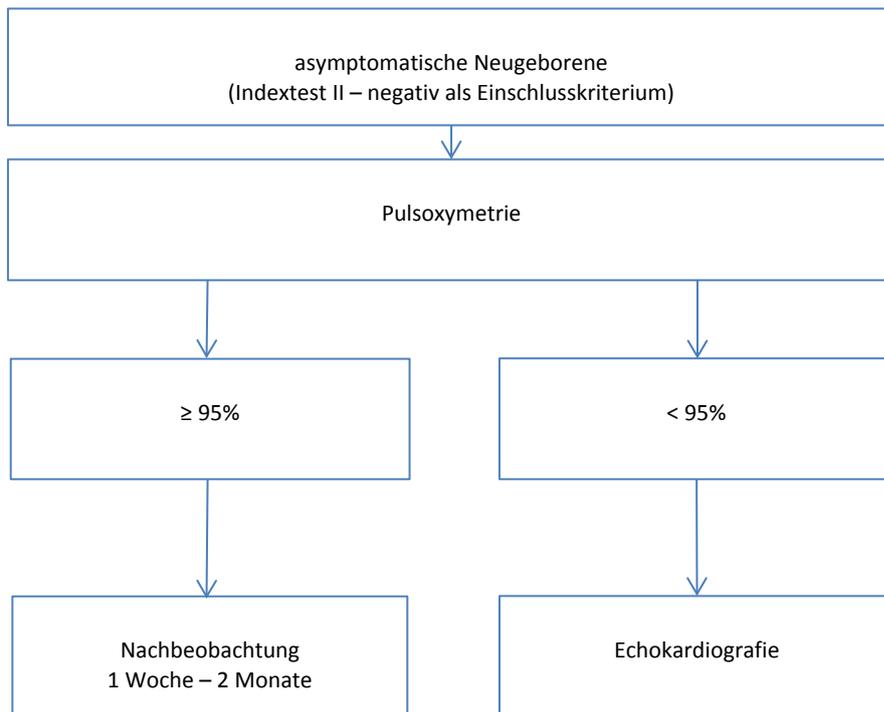


Abbildung 6: Screeningstrategie Ruangritnamchai 2007

Tautz 2010

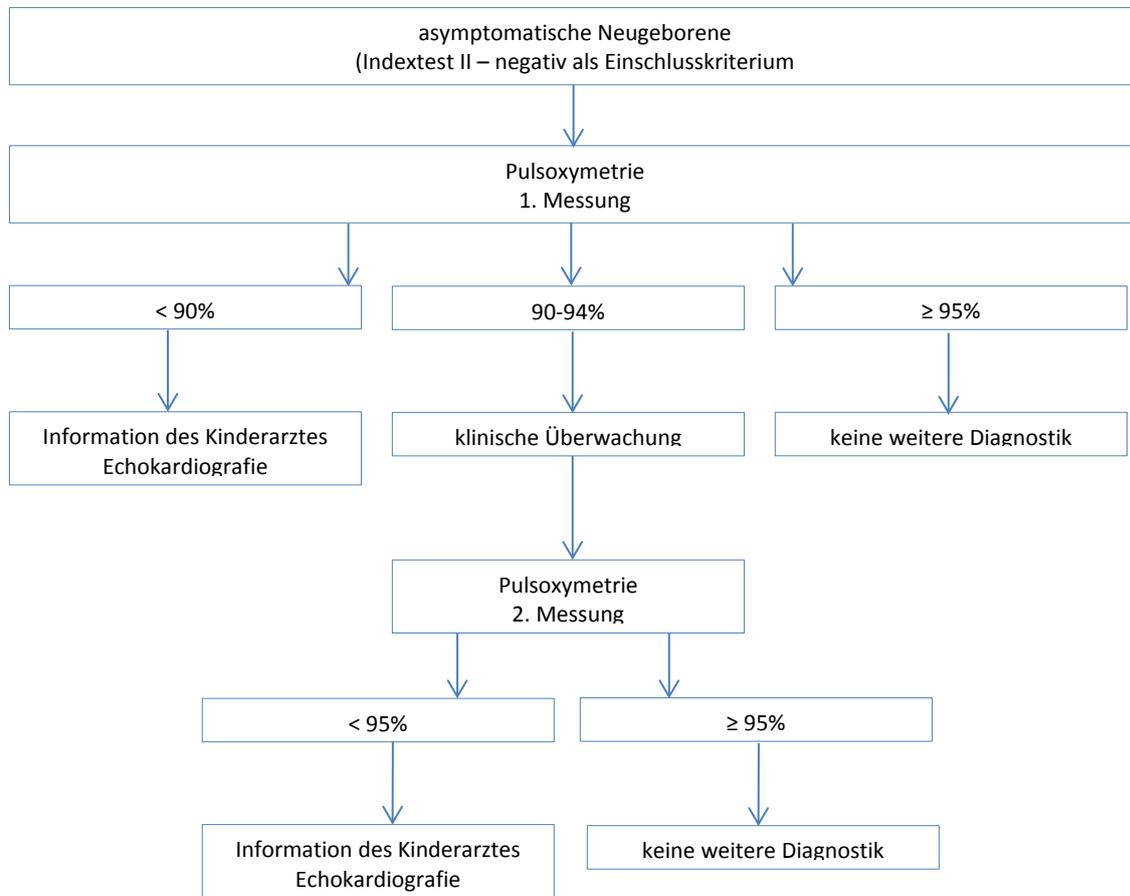


Abbildung 7: Screeningstrategie Tautz 2010

Zhao 2014

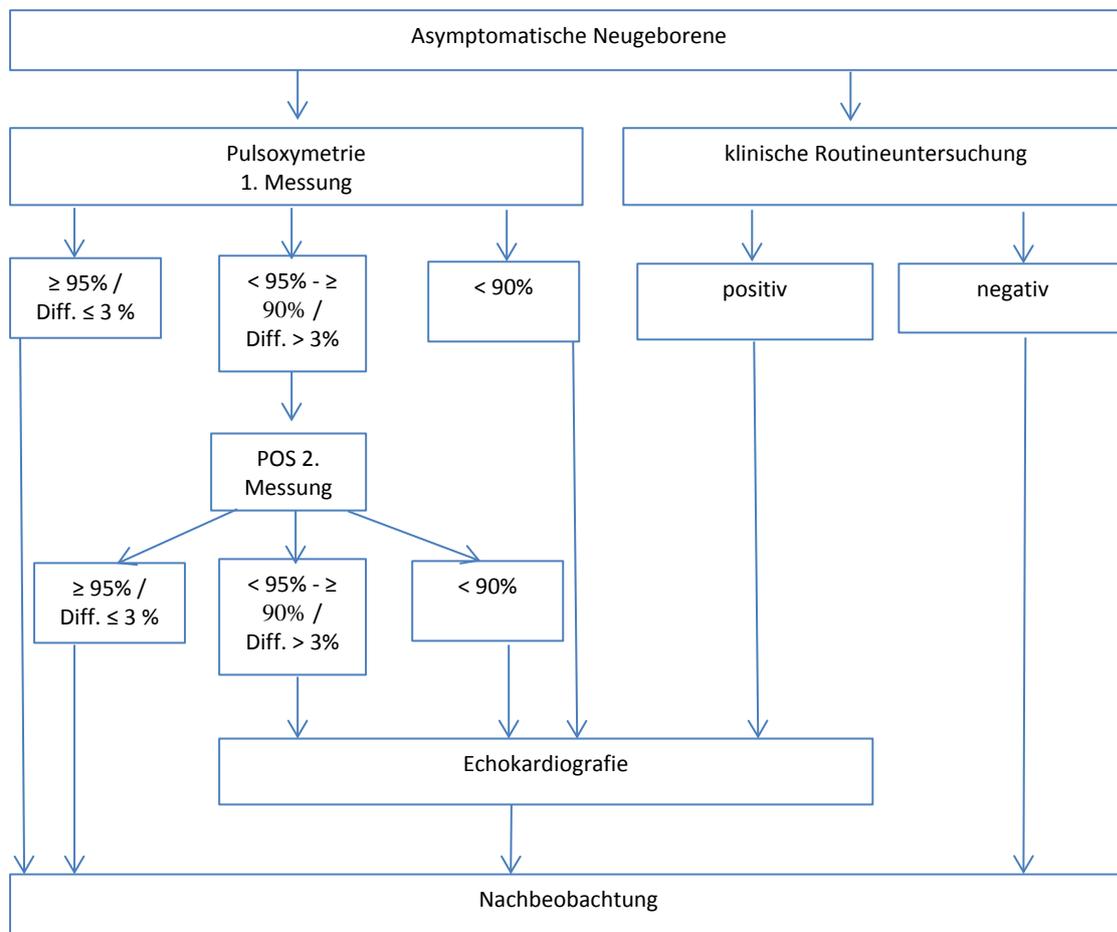


Abbildung 8: Screeningstrategie Zhao 2014