

IQWiG-Berichte – Nr. 463

Screening auf schwere kombinierte Immundefekte (SCID-Screening) bei Neugeborenen

Abschlussbericht

Auftrag: S15-02
Version: 1.0
Stand: 16.11.2016

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Screening auf schwere kombinierte Immundefekte (SCID-Screening) bei Neugeborenen

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

30.07.2015

Interne Auftragsnummer:

S15-02

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 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. Die Selbstangaben der externen Sachverständigen zu potenziellen Interessenkonflikten sind in Kapitel A8 dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Externe Sachverständige

- Michael Borte, Klinik für Kinder- und Jugendmedizin, Klinikum St. Georg gGmbH, Leipzig
- Tim Mathes, Monika Becker, Dawid Pieper, Institut für Forschung in der Operativen Medizin, Universität Witten / Herdecke, Köln

Das IQWiG dankt den extern Beteiligten für ihre Mitarbeit am Projekt.

Mitarbeiter des IQWiG¹

- Anne Rummer
- Elke Hausner
- Lina Rodenhäuser
- Stefan Sauerland
- Wiebke Sieben

Schlagwörter: Neugeborenen-Screening, Transplantation, schwerer kombinierter Immundefekt, Nutzenbewertung, Systematische Übersicht

Keywords: Neonatal Screening, Transplantation, Severe Combined Immunodeficiency, Benefit Assessment, Systematic Review

¹ Aufgrund gesetzlicher Datenschutzbestimmungen haben Mitarbeiter das Recht, ihrer Namensnennung nicht zuzustimmen.

Kernaussage

Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung eines Neugeborenen-screensings auf schwere kombinierte Immundefekte (SCID-Screening) in Kombination mit einer kurativen Therapie (allogenen Knochenmark- oder Stammzelltransplantation) im Vergleich zu keinem SCID-Screening hinsichtlich patientenrelevanter Endpunkte.

Fazit

Das Neugeborenencreening auf schwere kombinierte Immundefekte (SCID-Screening) in Kombination mit einer infektionsprophylaktischen Therapie, die in eine kurative Therapie mündet, zeigt einen Anhaltspunkt für einen Nutzen.

Für den Endpunkt Mortalität wird ein dramatischer Effekt zugunsten eines früheren versus einen späteren Beginn einer infektionsprophylaktischen Therapie beschrieben. Die Datenlage aus den Studien zur diagnostischen Güte reichte nicht aus, um die Sensitivität und Spezifität zu berechnen. Der positive prädiktive Wert (PPV) für ein SCID-Screening durch die Bestimmung von T-Cell Receptor Excision Circles (TREC) aus Filterkartenblut zeigt, dass der Test auf SCID grundsätzlich geeignet ist, Kinder mit SCID zu identifizieren. Unklar ist, wie viele Kinder nicht gefunden werden.

Diese Ergebnisse stützen sich auf 3 nicht randomisierte vergleichende Interventionsstudien zum Therapiebeginn sowie 5 Studien zur diagnostischen Güte, die mittels Linked-Evidence-Ansatz zusammengeführt wurden. Sämtliche Studien weisen ein hohes Verzerrungspotenzial auf.

Eine in Frankreich laufende vergleichende Interventionsstudie der Screeningkette, in der auch Daten zur Sensitivität und Spezifität erhoben werden sollen, lässt Ergebnisse zur Beantwortung der Nutzenfrage des SCID-Screenings erwarten; Daten werden bis Juni 2018 erhoben.

Inhaltsverzeichnis

	Seite
Kernaussage	iii
Tabellenverzeichnis	viii
Abbildungsverzeichnis	ix
Abkürzungsverzeichnis	x
1 Hintergrund	1
2 Fragestellung	3
3 Methoden	4
4 Ergebnisse	6
4.1 Ergebnisse der Informationsbeschaffung	6
4.2 Charakteristika der in die Bewertung eingeschlossenen Studien	6
4.2.1 Vergleichende Interventionsstudien der Screeningkette.....	6
4.2.2 Vergleichende Interventionsstudien zum Therapiebeginn.....	7
4.2.3 Studien zur diagnostischen Güte.....	8
4.3 Übersicht der vorhandenen bewertungsrelevanten Endpunkte / Zielgrößen ..	10
4.4 Bewertung des Verzerrungspotenzials auf Studien- und Endpunktebene	11
4.5 Ergebnisse zu patientenrelevanten Endpunkten	12
4.5.1 Ergebnisse zum Endpunkt Mortalität.....	12
4.5.2 Ergebnisse zum Endpunkt Morbidität (Infektionen).....	13
4.6 Ergebnisse zur diagnostischen Güte	13
4.7 Studien unklarer Relevanz	14
4.8 Landkarte der Beleglage	14
5 Einordnung des Arbeitsergebnisses	15
6 Fazit	18
Details des Berichts	19
A1 Projektverlauf	19
A1.1 Zeitlicher Verlauf des Projekts	19
A1.2 Dokumentation der Änderungen im Projektverlauf	20
A2 Details der Methoden	21
A2.1 Methodik gemäß Berichtsplan	21
A2.1.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette in die Untersuchung.....	21
A2.1.1.1 Population.....	21
A2.1.1.2 Prüf- und Vergleichsintervention.....	22

A2.1.1.3	Patientenrelevante Endpunkte.....	22
A2.1.1.4	Studientypen	22
A2.1.1.5	Studiendauer	23
A2.1.1.6	Tabellarische Darstellung der Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette).....	23
A2.1.2	Kriterien für den Einschluss von vergleichenden Interventionsstudien zum Therapiebeginn in die Untersuchung.....	23
A2.1.2.1	Population	24
A2.1.2.2	Prüf- und Vergleichsintervention.....	24
A2.1.2.3	Patientenrelevante Endpunkte.....	24
A2.1.2.4	Studientypen	24
A2.1.2.5	Studiendauer	24
A2.1.2.6	Tabellarische Darstellung der Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn).....	24
A2.1.3	Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Untersuchung.....	25
A2.1.3.1	Population	25
A2.1.3.2	Indextest.....	25
A2.1.3.3	Referenztest.....	25
A2.1.3.4	Zielgrößen	25
A2.1.3.5	Studientypen	26
A2.1.3.6	Studiendauer	26
A2.1.3.7	Tabellarische Darstellung der Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte).....	26
A2.1.4	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen.....	27
A2.1.5	Informationsbeschaffung	27
A2.1.5.1	Bibliografische Recherche	27
A2.1.5.2	Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien beziehungsweise Informationen zu relevanten Studien.....	28
A2.1.5.2.1	Systematische Übersichten.....	28
A2.1.5.2.2	Öffentlich zugängliche Studienregister	28
A2.1.5.2.3	Durch den G-BA übermittelte Unterlagen	28
A2.1.5.2.4	Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen.....	28
A2.1.5.2.5	Informationen aus Anhörungen.....	28
A2.1.5.3	Selektion relevanter Studien	29
A2.1.6	Informationsbewertung.....	29

A2.1.6.1	Bewertung von vergleichenden Interventionsstudien	29
A2.1.6.2	Bewertung von Studien zur diagnostischen Güte	31
A2.1.7	Informationssynthese und -analyse	31
A2.1.7.1	Gegenüberstellung der Ergebnisse der Einzelstudien	31
A2.1.7.2	Meta-Analysen	32
A2.1.7.2.1	Meta-Analysen für vergleichende Interventionsstudien.....	32
A2.1.7.2.2	Meta-Analysen für Studien zur diagnostischen Güte.....	32
A2.1.7.3	Aussagen zur Beleglage	33
A2.1.7.4	Sensitivitätsanalysen	34
A2.1.7.5	Subgruppenmerkmale und andere Effektmodifikatoren	34
A2.2	Spezifizierungen und Änderungen der Methodik.....	35
A3	Details der Ergebnisse	36
A3.1	Informationsbeschaffung.....	36
A3.1.1	Primäre Suchquellen.....	36
A3.1.1.1	Bibliografische Recherche	36
A3.1.1.2	Öffentlich zugängliche Studienregister.....	37
A3.1.2	Weitere Suchquellen.....	38
A3.1.2.1	Systematische Übersichten	39
A3.1.2.2	Durch den G-BA übermittelte Dokumente	39
A3.1.2.3	Anhörung	39
A3.1.2.4	Autorenanfragen.....	39
A3.1.3	Resultierender Studienpool.....	40
A3.1.4	Studien unklarer Relevanz.....	41
A3.2	Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudien der Screeningkette	42
A3.2.1	Studiendesign und Studienpopulationen	42
A3.2.2	Einschätzung des Verzerrungspotenzials auf Studienebene.....	43
A3.3	Patientenrelevante Endpunkte.....	43
A3.4	Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudien zum Therapiebeginn	44
A3.4.1	Studiendesign und Studienpopulationen – Studien zum Therapiebeginn	44
A3.4.2	Einschätzung des Verzerrungspotenzials auf Studienebene – Studien zum Therapiebeginn	48
A3.5	Patientenrelevante Endpunkte aus vergleichenden Interventionsstudien zum Therapiebeginn	49
A3.5.1	Verzerrungspotenzial auf Endpunktebene – Mortalität und Morbidität.....	50
A3.5.2	Ergebnisse zum Endpunkt Mortalität	52
A3.5.3	Ergebnisse zum Endpunkt Morbidität (Infektionen).....	53

A3.5.4	Subgruppenmerkmale und andere Effektmodifikatoren.....	54
A3.6	Charakteristika der in die Bewertung eingeschlossenen Studien zur diagnostischen Güte	55
A3.6.1	Studiendesign und Studienpopulationen	55
A3.6.2	Einschätzung des Verzerrungspotenzials	63
A3.6.2.1	Verzerrungspotenzial nach QUADAS 2.....	63
A3.6.2.2	Bedenken bezüglich der Übertragbarkeit nach QUADAS 2	65
A3.7	Ergebnisse zu den Zielgrößen	65
A3.7.1	Meta-Analysen.....	68
A3.7.2	Sensitivitätsanalysen.....	68
A3.7.3	Subgruppenmerkmale und andere Effektmodifikatoren.....	68
A4	Kommentare.....	69
A4.1	Bericht im Vergleich zu anderen systematischen Übersichten	69
A4.2	Kritische Reflexion des Vorgehens	70
A4.3	Würdigung der Stellungnahmen	71
A4.3.1	Würdigung des zusätzlichen Arguments: Gentherapie bei ADA-SCID	71
A4.3.2	Würdigung der zusätzlichen Argumente zum relevanten Studienpool	72
A4.3.2.1	Berücksichtigung von Kwan 2014 und Pai 2014.....	72
A4.3.2.2	Berücksichtigung von Grazioli 2014, la Marca 2014 und Speckmann 2012.....	73
A4.3.3	Würdigung des zusätzlichen Arguments: Ergebnis der Nutzenbewertung	73
A5	Literatur	74
A6	Studienlisten	81
A6.1	Liste der eingeschlossenen Studien.....	81
A6.2	Liste der gesichteten systematischen Übersichten	82
A6.3	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	82
A7	Suchstrategien	89
A7.1	Suchstrategien in bibliografischen Datenbanken.....	89
A7.1.1	Embase.....	89
A7.1.2	MEDLINE	90
A7.1.3	PubMed.....	91
A7.1.4	The Cochrane Library.....	92
A7.2	Suche in Studienregistern.....	93
A7.2.1	ClinicalTrials.gov	93
A7.2.2	EU Clinical Trials Register.....	93
A7.2.3	International Clinical Trials Registry Platform Search Portal.....	93
A8	Offenlegung potenzieller Interessenkonflikte	94
A8.1	Darlegung potenzieller Interessenkonflikte der externen Sachverständigen ...	94

Tabellenverzeichnis

	Seite
Tabelle 1: Matrix der Endpunkte	11
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)	23
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)	25
Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)	27
Tabelle 5: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	34
Tabelle 6: In Studienregistern identifizierte Studien unklarer Relevanz	38
Tabelle 7: Übersicht zu Autorenanfragen	40
Tabelle 8: Studienpool	41
Tabelle 9: Charakterisierung der eingeschlossenen Studie – Studie zur Screeningkette	42
Tabelle 10: Ein- / Ausschlusskriterien für Personen in der Studie – Studie zur Screeningkette	43
Tabelle 11: Charakterisierung der Studienpopulationen – Studie zur Screeningkette	43
Tabelle 12: Charakterisierung der Intervention	43
Tabelle 13: Charakterisierung der eingeschlossenen Studien – Studien zum Therapiebeginn	45
Tabelle 14: Ein- / Ausschlusskriterien für Patienten – Studien zum Therapiebeginn	46
Tabelle 15: Charakterisierung der Studienpopulationen – Studien zum Therapiebeginn	47
Tabelle 16: Charakterisierung der Interventionen – Studien zum Therapiebeginn	48
Tabelle 17: Verzerrungspotenzial auf Studienebene – Studien zum Therapiebeginn	49
Tabelle 18: Übersicht zur Extraktion von patientenrelevanten Endpunkten, Datenverfügbarkeit	50
Tabelle 19: Verzerrungspotenzial auf Endpunktebene – Mortalität	51
Tabelle 20: Verzerrungspotenzial auf Endpunktebene – Morbidität	51
Tabelle 21: Ergebnisse – Mortalität	53
Tabelle 22: Ergebnisse – Morbidität	54
Tabelle 23: Charakterisierung der eingeschlossenen Studien – Studien zur Testgüte	56
Tabelle 24: Ein- / Ausschlusskriterien für Personen in den Studien – Studien zur Testgüte	58
Tabelle 25: Charakterisierung der Studienpopulationen – Studien zur Testgüte	58
Tabelle 26: Charakterisierung von Index- und Referenztest	59
Tabelle 27: Verzerrungspotenzial nach QUADAS 2	64
Tabelle 28: Bedenken bezüglich der Übertragbarkeit nach QUADAS 2	65
Tabelle 29: Ergebnisse der Studien zur diagnostischen Güte des Screeningtests auf SCID	67

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion.....	37
Abbildung 2: Gepoolter Gesamteffekt Therapiebeginn	53

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ADA	Adenosindeaminase
ELVIS	Erreger, Lokalisation, Verlauf, Intensität, Summe (Akronym)
EMA	European Medicines Agency (Europäische Zulassungsbehörde)
GARFIELD	Granulome, Autoimmunität, rezidivierendes Fieber, ungewöhnliche Ekzeme, Lymphoproliferation, chronische Darmentzündung (Akronym)
G-BA	Gemeinsamer Bundesausschuss
GVHD	Graft-versus-Host-Disease
ICD	International Statistical Classification of Diseases and Related Health Problems (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KREC	Kappa-deleting Recombination Excision Circles
PCR	Polymerase Chain Reaction (Polymerase-Kettenreaktion)
PNP	Purinnucleosidphosphorylase
PPV	Positive predictive Value (positiver prädiktiver Wert)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
SCID	Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)
TREC	T-Cell Receptor Excision Circles
VOPT	Verification of only positive Testers
ZAP70	Zeta-Chain-associated Protein 70 kDa

1 Hintergrund

Der schwere kombinierte Immundefekt (Severe combined Immunodeficiency, SCID) beschreibt eine Gruppe genetischer Erkrankungen, die durch ein völliges Fehlen der Immunabwehr charakterisiert sind. Wie viele Kinder in Deutschland mit SCID geboren werden, ist nicht bekannt. Statistiken der gesetzlichen Krankenversicherung berichten für 2013 von 21 Fällen bei unter 1-jährigen Kindern [1]. Hierzu passen Daten aus USA und Großbritannien, wo eine Inzidenz von 1:30 000 bis 58 000 errechnet wurde [2-4].

Die Störung der Immunfunktionen resultiert aus einer Hemmung der Entwicklung von für die Abwehr von Infektionen erforderlichen T-Lymphozyten (T-Zellen), je nach zugrunde liegendem genetischen Defekt begleitet von einem Fehlen auch der B-Lymphozyten (B-Zellen) sowie der NK-Lymphozyten (NK-Zellen). International erfolgt die Einteilung nach Immunphänotyp in die Gruppen T^+B^- und T^-B^- [5]. In der ICD-10 sind unter der Gruppe der SCID retikuläre Dysgenese und Adenosindeaminase(ADA)-Mangel (beide T^+B^-) als eigene Ziffern angeführt. Eine weitere Gruppe bildet hier der Purinnucleosidphosphorylase(PNP)-Mangel (T^+B^{variabel}) [6]. Andere Einteilungen folgen pathogenetischen Kriterien, also nach zugrunde liegendem Gendefekt [7].

Die klinische Manifestation ist von der Erregerexposition abhängig. Kinder mit SCID weisen aufgrund der fehlenden Immunabwehr bereits im Säuglingsalter eine pathologische Infektanfälligkeit sowie Infektionskomplikationen mit hoher Neigung zu opportunistischen und rekurrenden Infektionen auf. Lebendimpfungen wie zum Beispiel die in der Schutzimpfungs-Richtlinie des G-BA vorgesehene Rotavirus-Impfung [8] sind bei diesen Kindern kontraindiziert [7,9], Gleiches gilt für die Transfusion unbestrahlter Blutprodukte [7,10,11].

In der Leitlinie „Diagnostik von primären Immundefekten“ haben die Verfasser die klinischen Kriterien für die Diagnose primärer Immundefekte, die auch SCID umfassen, anhand der Akronyme ELVIS (für: Erreger, Lokalisation, Verlauf, Intensität und Summe) und GARFIELD (für: Granulome, Autoimmunität, rezidivierendes Fieber, ungewöhnliche Ekzeme, Lymphoproliferation und chronische Darmentzündung) gelistet [12], weiter werden Gedeihstörung sowie eine auffällige Familienanamnese und Konsanguinität genannt [7,12,13]. Der Immunstatus gibt Aufschluss über die zugrunde liegende Erkrankung. Dieser wird mittels einer speziellen immunologischen Labordiagnostik bestimmt, und zwar durch Lymphozytenphänotypisierung (auch: Berechnung der absoluten Lymphozytenzahlen, Lymphocyte counts); ferner wird ein antikörperbasiertes Nachweisverfahren beschrieben, Enzyme-linked Immunosorbent Assay [4,14]. Charakteristisch für SCID ist eine T-Zell-Lymphopenie [13]; als Grenzwert werden Werte zwischen 25 und 252 Zellen/ μ l abhängig von der angewandten Methode angegeben [15-18]. Bei Verdacht auf ADA- und PNP-Mangel kann die Diagnostik ferner über Enzymbestimmungen erfolgen [4,14]. Darüber hinaus sind molekulargenetische Untersuchungen Bestandteil der Diagnostik [7,10,12].

Unbehandelt sterben die meisten Kinder mit SCID innerhalb von 1 bis 2 Jahren [4,7,9,10].

Etablierte kurative Therapie ist die allogene Knochenmark- [7] oder Stammzelltransplantation zur Rekonstitution einer intakten Immunfunktion [4]. Bei ADA- und PNP-Mangel kommen als Behandlungsalternativen grundsätzlich eine Enzyersatz- und Gentherapie in Betracht, wobei Erstere nicht kurativ ist und Letztere sich noch im Entwicklungsstadium befindet ([19]; zur Zulassung eines Gentherapeutikums zur Behandlung von ADA-SCID siehe [20]). Bereits vor Einleitung einer kurativen Therapie müssen die Kinder durch präventive und supportive Maßnahmen stabilisiert werden; hierzu zählen strenge hygienische Vorsichtsmaßnahmen, Prophylaxe opportunistischer Infektionen sowie weitere Maßnahmen wie die Supplementation von Immunglobulinen, Bestrahlung von Bluttransfusionen oder der Verzicht auf Lebendimpfungen [7,9,10].

In Deutschland werden Neugeborene entsprechend dem erweiterten Neugeborenencreening der Kinder-Richtlinien des G-BA [21] auf die dort genannten Zielkrankheiten untersucht. Dazu wird im Alter von 48 bis 72 Lebensstunden Filterkartenblut gewonnen (Probe aus nativem Venen- oder Fersenblut, das auf Filterpapierkarten aufgetropft und getrocknet wird). Diese Probe könnte ebenfalls zur Bestimmung der T- bzw. T- und B-Zell-Defizienz verwendet werden. Mittels Polymerase-Kettenreaktion (Polymerase Chain Reaction, PCR) könnten hieraus T-Cell Receptor Excision Circles (TREC) und bei Bestimmung der B-Zell-Defizienz Kappa-deleting Recombination Excision Circles (KREC) quantifiziert werden. TREC und KREC sind kleine ringförmige DNA-Fragmente, die bei der Biosynthese des T- beziehungsweise B-Zell-Rezeptors als Abfallprodukte entstehen und deren Anzahl im Blut mit der Zahl frisch entstehender naiver T- beziehungsweise B-Zellen korreliert. Darüber hinaus ist es möglich, für den methodischen Erfolg der Messung von TREC- und KREC-Kopien Beta-Actin als Kontrollmarker zu quantifizieren [4,9].

Seit 2010 wird in den USA auf SCID gescreent, wobei dort das Blut allein auf TREC untersucht wird [3,18]. Eine zusätzliche Bestimmung von KREC dient der Entdeckung bestimmter SCID-Erkrankungen, die allein durch TREC nicht auffallen, beispielsweise in Hinblick auf den ADA-Mangel [9]. Grundsätzlich erlaubt das Screening auf SCID auch die Identifizierung von Neugeborenen mit anderen Immundefizienzen als SCID [9].

Ziel eines Neugeborenencreenings auf SCID ist die Identifikation von betroffenen Kindern vor der ersten Infektion und die frühzeitige Einleitung einer idealerweise kurativen Therapie.

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung eines Neugeborenen-screensings auf schwere kombinierte Immundefekte (SCID-Screening) in Kombination mit einer kurativen Therapie (allogenen Knochenmark- oder Stammzelltransplantation) im Vergleich zu keinem SCID-Screening hinsichtlich patientenrelevanter Endpunkte.

3 Methoden

In die Nutzenbewertung wurden vergleichende Interventionsstudien der Screeningkette eingeschlossen. Für den Fall, dass solche Studien nicht oder in nicht ausreichender Qualität vorliegen sollten, war eine Bewertung vergleichender Interventionsstudien zum Therapiebeginn sowie von Studien zur diagnostischen Güte als die einzelnen Bausteine der Screeningkette vorgesehen (Linked Evidence).

Vergleichende Interventionsstudien der Screeningkette

Die Zielpopulation der Nutzenbewertung anhand vergleichender Interventionsstudien der Screeningkette bildeten Neugeborene. Die Prüflintervention bildete das SCID-Screening bei Neugeborenen unter Verwendung von Filterpapierkarten. Als Vergleichsintervention galt die Neugeborenenversorgung ohne SCID-Screening.

Für die Untersuchung wurden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtüberleben, krankheitsspezifisches Überleben),
- Morbidität (z. B. Infektionen oder Organschäden, Transplantationen),
- Krankenhausaufenthalte (insbesondere auch solche in Umkehrisolierung),
- Entwicklungsstörungen (z. B. Störungen der kognitiven, psychosozialen, emotionalen, grob- und feinmotorischen Entwicklung),
- unerwünschte Ereignisse,
- gesundheitsbezogene Lebensqualität des Kindes (gemessen z. B. durch Proxy Rating).

Es sollten randomisierte kontrollierte Studien (RCTs) in die Nutzenbewertung eingeschlossen werden. Sofern die auf RCTs basierende Datenlage zur Nutzenbewertung nicht ausreichte, wurden auch vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) eingeschlossen. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Vergleichende Interventionsstudien zum Therapiebeginn

Es wurde der Nutzen eines früheren gegenüber dem eines späteren Therapiebeginns erfasst. Eingeschlossen wurden vergleichende Interventionsstudien zum Therapiebeginn. Die Zielpopulation bildeten Neugeborene mit SCID, deren Diagnosestellung übertragbar sein musste auf die einer Screeningsituation. Als Vergleichsintervention galt ein späterer Therapiebeginn. Es wurden die oben unter „Vergleichende Interventionsstudien der Screeningkette“ genannten patientenrelevanten Endpunkte betrachtet. Neben RCTs wurden vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) eingeschlossen. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Studien zur diagnostischen Güte

In die Bewertung wurden Studien mit Neugeborenen eingeschlossen. Der Indextest war die Testung auf SCID unter Verwendung von Filterpapierkarten. Den Referenztest bildeten die Lymphozytenphänotypisierung beziehungsweise Enzymbestimmung und/oder die genetische Analyse sowie, bei unauffälligem Befund, zusätzlich die Nachbeobachtung. Eingeschlossen wurden Studien, aus denen personenbezogene Daten zur Berechnung der diagnostischen Güte in Hinblick auf die Entdeckung von SCID ableitbar waren.

Literaturrecherche

Eine systematische Literaturrecherche nach Primärliteratur wurde in den Datenbanken MEDLINE, Embase und Cochrane Central Register of Controlled Trials durchgeführt. Parallel erfolgte eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE, Embase, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects und Health Technology Assessment Database.

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

Die Selektion relevanter Studien erfolgte von 2 Reviewern unabhängig voneinander. Die Datenextraktion erfolgte in standardisierte Tabellen. Zur Einschätzung der qualitativen Ergebnissicherheit wurde das Verzerrungspotenzial auf Studien- und Endpunktebene bewertet und jeweils in niedrig oder hoch eingestuft. Die Ergebnisse der einzelnen Studien wurden nach Endpunkten geordnet beschrieben.

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren und keine bedeutsame Heterogenität beobachtet wurde, wurden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst.

Für jeden Endpunkt wird eine Aussage zur Beleglage des (Zusatz-)Nutzens und (höheren) Schadens in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (Zusatz-)Nutzen oder (höheren) Schaden vor“ getroffen.

4 Ergebnisse

4.1 Ergebnisse der Informationsbeschaffung

Die systematische Literaturrecherche in den bibliografischen Datenbanken ergab 12 Publikationen zu 12 Studien, die die für diesen Bericht definierten Kriterien zum Studieneinschluss erfüllten. Die letzte Suche fand am 15.08.2016 statt.

Durch die Suche in den weiteren Suchquellen wurden keine zusätzlichen relevanten Dokumente identifiziert. Informationen aus Autorenanfragen flossen in die Bewertung ein.

Es wurde 1 Studie identifiziert, deren Relevanz nicht abschließend geklärt werden konnte. Des Weiteren wurden 5 laufende Studien identifiziert.

Insgesamt wurden somit 12 Studien (12 Dokumente) als relevant für die Fragestellung der vorliegenden Nutzenbewertung identifiziert.

4.2 Charakteristika der in die Bewertung eingeschlossenen Studien

Für die Nutzenbewertung wurden vergleichende Interventionsstudien der Screeningkette herangezogen. Für den Fall, dass diese nicht in ausreichender Anzahl oder Qualität vorlagen, waren sodann vergleichende Interventionsstudien zum Therapiebeginn sowie Studien zur diagnostischen Güte heranzuziehen.

4.2.1 Vergleichende Interventionsstudien der Screeningkette

Als einzige vergleichende Interventionsstudie der Screeningkette wurde die retrospektive Analyse zur Bestimmung der Kosteneffektivität eines Neugeborenen Screenings auf SCID von **Kubiak 2014** [22] eingeschlossen. Die Studie verglich zwischen Behandlungszentren, die in US-amerikanischen Bundesstaaten mit oder ohne SCID-Screening liegen. In den Analysen betrachteten die Autoren jedoch den Vergleich zwischen einer früheren und späteren Behandlung ($< 3,5$ versus $\geq 3,5$ Monate Lebensalter). Wegen der Angabe zur Art der Diagnose (Screening oder kein Screening) und der Möglichkeit, dies dem Therapiezeitpunkt zuzuordnen, konnte diese Studie als vergleichende Interventionsstudie der Screeningkette eingeschlossen werden.

Die Ergebnisse der Studie sind allerdings nicht interpretierbar und werden in diesem Bericht nicht dargestellt. Es ist nicht berichtet, zu welchem Zeitpunkt innerhalb der 6-jährigen Studienlaufzeit (2005 bis 2011) die Neugeborenen eingeschlossen wurden. Für die mittels Screening identifizierten Kinder (Interventionsgruppe) ergibt sich dies aus dem Zeitpunkt der Einführung des Screenings im August 2010 [23]. So ist davon auszugehen, dass sie erst gegen Ende des Studienzeitraums gefunden wurden und ihre Nachbeobachtungszeit deutlich kürzer sein kann als die der Babys in der Vergleichsgruppe. Somit kann der Gruppenunterschied in den berichteten Endpunkten nicht auf einen Zeitraum bezogen werden und damit ist die Vergleichbarkeit der Nachbeobachtungsdauer in den Gruppen unklar. Ferner können die

Ergebnisse für die berichteten Endpunkte (abgesehen von der Mortalität) nicht eindeutig den Gruppen Screening / kein Screening zugeordnet werden.

4.2.2 Vergleichende Interventionsstudien zum Therapiebeginn

Als vergleichende Interventionsstudien zum Therapiebeginn wurden 3 Studien eingeschlossen.

In der vergleichenden retrospektiven Datenanalyse **Brown 2011** [24] waren 108 Kinder aus 2 Krankenhäusern in England (UK) mit SCID eingeschlossen, die eine Stammzelltransplantation erhalten haben. Die Gruppe mit früherem Therapiebeginn bildeten Kinder, die aufgrund einer positiven Familienhistorie, meistens aufgrund eines betroffenen Geschwisterkindes, pränatal oder bei der Geburt getestet wurden und bei denen SCID diagnostiziert wurde (n = 60). Die betroffenen Geschwisterkinder (oder andere Verwandte) bildeten die Vergleichsgruppe der späteren Intervention. Ihre Diagnose war erst nach auffälliger Symptomatik erfolgt (n = 48). Damit ging eine Selektion der Studienteilnehmer einher, da so allein bis dahin überlebende Kinder eingeschlossen wurden. Der Therapiezeitpunkt selbst ist nicht berichtet. Die Nachbeobachtungsdauer ist ebenfalls nicht berichtet. In dieser Publikation werden die Ergebnisse zu den Endpunkten Mortalität (vor und nach Transplantation) und Morbidität berichtet, wobei für die Morbidität (Infektionen) in der Vergleichsgruppe für einen Großteil der Kinder (n = 19) Angaben fehlen.

In der retrospektiven / prospektiven Studie **Myers 2002** [25] wird von 117 Kindern mit Stammzelltransplantation aufgrund diagnostiziertem SCID berichtet. Das Hauptaugenmerk der Studie liegt auf dem Vergleich der Entwicklung des Immunsystems bei Kindern, die vor beziehungsweise nach dem 28. Lebenstag transplantiert wurden. Alle Kinder der Gruppe früherer Therapiebeginn (vor dem 28. Lebenstag, n = 21) waren aufgrund einer positiven Familienhistorie pränatal oder bei der Geburt diagnostiziert worden; der Diagnosezeitpunkt der Gruppe mit späterem Therapiebeginn (nach dem 28. Lebenstag, n = 96) ist nicht berichtet. Die Nachbeobachtungsdauer ist auch in dieser Studie nicht berichtet. Da die Autoren in ihre Auswertung nur die Überlebenden der Vergleichsgruppe eingeschlossen haben, werden vorliegend nur die Daten für den Endpunkt Mortalität berichtet.

In der vergleichenden retrospektiven Kohortenstudie **Cuvelier 2016** [26] konnten für den interessierenden Vergleich Daten von 8 Kindern herangezogen werden. Es handelte sich um Kinder aus 6 mennonitischen Familien jeweils mit Zeta-Chain-associated Protein 70 kDa Deficiency (ZAP70-Defizienz, SCID im Sinne des Berichts), von denen 2 aufgrund positiver Familienhistorie asymptomatisch identifiziert und innerhalb der ersten beiden Lebensmonate transplantiert wurden. Als patientenrelevante Endpunkte werden Daten zur Mortalität sowie Morbidität (Infektionen zum Zeitpunkt der Transplantation) berichtet. Da die Autoren das einzige verstorbene Kind aus der Analyse ausgeschlossen haben, bei dem das Vorliegen eines SCID nicht abschließend diagnostiziert wurde, und darüber hinaus eine Zuordnung zu einer der Gruppen nicht möglich ist, waren die Daten zur Mortalität nicht interpretierbar. In der Studie werden weitere Endpunkte zur Morbidität (zum Beispiel Graft-versus-Host-Disease

[GVHD], Anzahl Transplantationen) berichtet. Da nur insgesamt 8 Kinder in dieser Studie untersucht wurden, kann sie in Bezug auf Endpunkte, zu denen keine weiteren Studien Daten berichtet haben, keine belastbare Datengrundlage für die Nutzenbewertung beitragen. Auf die Darstellung wird daher verzichtet.

4.2.3 Studien zur diagnostischen Güte

Es wurden 8 Studien zur diagnostischen Güte eingeschlossen [17,23,27-32]. Bei allen Studien handelt es sich um prospektive Kohortenstudien, in denen ausschließlich die positiven Testergebnisse verifiziert wurden (VOPT). Die Ergebnisse von 3 dieser Studien [17,28,29] werden in diesem Bericht wie nachfolgend erläutert nicht dargestellt.

Von den 8 Studien berichten 5 die Ergebnisse des um ein SCID-Screening erweiterten Neugeborenencreenings in einzelnen Staaten der USA bzw. der Navajo-Nation [17,23,27-29]. Die Screeningalgorithmen unterscheiden sich im Detail in Hinblick auf die Trennwerte sowie auf den Einsatz von Kontrollmarkern zur Prüfung der Probenqualität. Übereinstimmend wurden aber jeweils Proben aus beim Neugeborenencreening gewonnenem Filterkartenblut genommen, um diese auf die Anzahl von TRECs als Marker für SCID zu testen.

Kwan 2015 beschreibt die Ergebnisse des SCID-Screenings von Februar 2012 bis Juli 2014 in Navajo Nation, USA, einer Population, die sich durch eine sehr hohe Inzidenz von SCIDs auszeichnet [27]. Ergebnisse aus dem vorangehenden Zeitraum seit März 2009 beziehen sich auf die vorgeschaltete Pilotstudie, die in der vorliegenden Auswertung aufgrund des explorativen Charakters keine Berücksichtigung finden. Im Rahmen der Anwendungsstudie wurde das Filterkartenblut von 6100 Neugeborenen getestet. Der Screeningalgorithmus sah eine 2-stufige Messung bei auffälligem ersten Testergebnis vor (Trennwert der 1. Messung: 40 TRECs, Trennwert der 2. Messung: 25 TRECs). Neugeborene ohne nachweisbare TRECs wurden als „testpositiv“ oder mit einem Nachweis von 1 bis 25 TRECs als „vermutlich testpositiv“ sofort dem Referenztest zugewiesen, andernfalls folgte eine Wiederholungsmessung. Auf ein abschließend auffälliges Testergebnis folgte als Referenztest neben einer klinischen Untersuchung die Erstellung eines immunologischen Profils.

In einer früheren Studie berichtet **Kwan 2013** [23] die Ergebnisse der ersten beiden Jahre nach der Einführung des SCID-Screenings in Kalifornien, USA. Von August 2010 bis August 2012 durchliefen 993 724 Neugeborene das Screening. Wie beim Screening an der Navajo-Population sah der Screeningalgorithmus eine 2-stufige Messung vor. Abweichungen gab es bei der Interpretation der Testergebnisse, indem Testergebnisse mit ≤ 5 TRECs bereits als positiv galten. Abschließend positive Testergebnisse wurden anhand einer Vollblutanalyse (großes Blutbild) mittels Durchflusszytometrie oder einer Lymphozytensubgruppenbestimmung verifiziert.

Vogel 2014 [28] beschreibt die Ergebnisse der ersten beiden Jahre des 2010 in New York State, USA eingeführten SCID-Screenings (September 2010 bis September 2012). In dieser Zeit durchliefen 485 912 Neugeborene das Screening. Der Fokus der Publikation liegt auf der

Beschreibung der Methode zur Probenanalyse sowie des Follow-up-Prozesses. Dazu gehörte die Modifizierung des ursprünglich auf ≤ 200 TRECs (Durchschnittswert von 3 Messungen) angesetzten Trennwertes: Während der Studienlaufzeit fügte die Forschergruppe zur Optimierung der Testergebnisse zunächst eine sogenannte *borderline*-Kategorie hinzu (150–200 TRECs), die dann nochmals angepasst wurde (125–200 TRECs). Als Referenztest diente eine Vollblutanalyse mittels Durchflusszytometrie; abweichend konnte die Zellaktivierung mittels Mitogentest bestimmt, eine Chromosomenanalyse oder ein Gentest durchgeführt werden. Um die Studienergebnisse trotz der Änderungen des Trennwertes zur Bestimmung der diagnostischen Güte des Tests verwenden zu können, sollten die Ergebnisse für einen konstanten Trennwert von ≤ 200 TRECs herangezogen werden. Allerdings wurde in der Studie nur etwa die Hälfte der positiven Testergebnisse und damit für eine Bewertung der diagnostischen Güte nicht ausreichend viele in diesem Sinne auffällige Testergebnisse verifiziert. Die Studienergebnisse können daher nicht zur Beschreibung der diagnostischen Güte verwendet werden und werden vorliegend nicht dargestellt.

Die Studie **Routes 2009** [17] berichtet die Ergebnisse des ersten Jahres nach der Einführung des SCID-Screenings in Wisconsin, USA von Januar bis Dezember 2008, in dem 71 000 Neugeborene gescreent wurden. Auch in Wisconsin wurden Proben aus Filterkartenblut auf die Anzahl von TRECs als Marker für SCID getestet. Bei Unterschreiten des Trennwertes von 25 TRECs war eine 2. Messung vorgesehen. Abschließend auffällige Messergebnisse waren anhand einer Lymphozytensubgruppenanalyse mittels Durchflusszytometrie als Referenztest zu überprüfen. Dieser Referenztest wurde allerdings nur unvollständig angewendet. Ursache hierfür ist bei Berücksichtigung sowohl früh- als auch termingeborener Kinder in erster Linie die hohe Anzahl von Todesfällen ohne Angabe der Todesursache, bei der das positive Screeningergebnis nicht mittels Referenztest bestätigt wurde: Insgesamt haben nur 11 von 42 positiv getesteten Neugeborenen den Referenztest erhalten. Damit wurden weniger als 70 % der positiven Screeningergebnisse verifiziert. Daher sind die Daten nicht in die Nutzenbewertung eingeflossen.

Die Studie **Verbsky 2012** [29] berichtet von den ersten 3 Jahren nach der Einführung des Screenings in Wisconsin, USA von Januar 2008 bis Dezember 2011. Die für die vorliegende Bewertung nicht verwertbaren Ergebnisse von Routes 2009 [17] (siehe oben) sind hiervon umfasst; insgesamt haben im 2 Jahre längeren Studienzeitraum 207 696 Neugeborene das SCID-Screening im Rahmen des Neugeborenencreenings durchlaufen. Nach einer Laufzeit von 19 Monaten wurde der ursprüngliche Trennwert von < 25 TRECs aufgrund der geringen Anzahl identifizierter Kinder mit einem als *abnormal* klassifizierten Testergebnis auf < 40 TRECs angehoben. Eine Ergebnisdarstellung, die eine Zuordnung der Testergebnisse zu den verwendeten Trennwerten erlaubt, fehlt in der Publikation ebenso wie eine Darstellung des Patientenflusses inklusive der Angabe und Erläuterung der Anzahl der durchgeführten sowie nicht durchgeführten Referenztests. Aus diesem Grund sind auch die Ergebnisse dieser Studie für die Bewertung der diagnostischen Güte des SCID-Screenings nicht verwertbar und werden nicht dargestellt.

Chien 2015 [30] berichtet von dem in Taiwan, im University Hospital Newborn Screening Center, in das reguläre Neugeborenen-Screening integrierten SCID-Screening. Zwischen Mai 2010 und Dezember 2011 wurden 106 391 Neugeborene unter Verwendung von Proben aus Filterkartenblut durch eine Bestimmung der Anzahl von TRECs auf SCID getestet. Der Trennwert war bei der 2-stufigen Testung auf < 40 TRECs festgelegt. Bei Unterschreiten des Trennwertes (zwischen > 0 und 40 TRECs) galt das Testergebnis als *inconclusive*; sofern der Wiederholungstest das Ergebnis bestätigte, galt der Test als positiv. Abweichend hiervon wurde bei 0 TRECs bereits in der ersten Messung von einem positiven Testergebnis ausgegangen. Der Referenztest bestand bei positivem Testergebnis aus einer Vollblutanalyse (großes Blutbild mittels Durchflusszytometrie) sowie einem Test auf das Mikrodeletionssyndrom 22q11.2, mit dem der Nachweis beispielsweise des DiGeorge-Syndroms möglich ist, aber auch anderer Krankheiten, die nicht unter SCID fallen (sogenannter TUPLE1-Test). Neugeborene mit Testergebnissen, die als *inconclusive* gewertet wurden, erhielten nur den TUPLE1-Test.

Kanegae 2016 [31] beschreibt die Ergebnisse eines SCID-Screenings von 8682 Neugeborenen in 2 Krankenhäusern in Brasilien, wobei Angaben zum Zeitraum fehlen. Filterkartenblut von Neugeborenen wurde auf TRECs getestet. Der Testalgorithmus sah auch hier eine 2-stufige Messung vor, wobei der Trennwert bei < 30 TRECS lag. Ein abschließend positives Testergebnis wurde mittels Vollblutanalyse sowie Lymphozytensubgruppenanalyse auf T-, B- und NK-Zellen verifiziert.

Als einzige der eingeschlossenen Studien berichtet **de Felipe 2016** [32] von einem kombinierten SCID-Screening, bei dem anhand von Proben aus Filterkartenblut aus dem regulären Neugeborenen-Screening neben TRECs auch KRECs quantifiziert wurden. Berichtet werden Daten von 5160 Neugeborenen von Februar 2014 bis März 2015 in Sevilla, Spanien. Als Trennwerte sind < 6 TRECs und / oder < 4 KRECs pro Probe angegeben. Der 3-stufige Screeningalgorithmus sah bei Unterschreiten der Trennwerte Wiederholungsmessungen vor; abschließend positive Testergebnisse wurden anhand einer Lymphozytensubgruppenanalyse mittels Durchflusszytometrie überprüft. Die Ergebnisse dieser Studie umfassen die Ergebnisse einer dazugehörigen Pilotstudie [33], in der der Trennwert mehrfach angepasst wurde. Mit Entwicklung und Überprüfung des Trennwertes anhand eines sich überschneidenden Samples sind zur Bewertung der diagnostischen Güte nur die um die Daten aus der Pilotstudie bereinigten Ergebnisse aus de Felipe 2016 verwertbar.

4.3 Übersicht der vorhandenen bewertungsrelevanten Endpunkte / Zielgrößen

Aus **Kubiak 2014** [22] als **vergleichende Interventionsstudie der Screeningkette** konnten keine Daten zu patientenrelevanten Endpunkten verwertet werden.

Aus 3 **Studien zum Vergleich früherer versus späteren Therapiebeginn** konnten Daten zu patientenrelevanten Endpunkten extrahiert werden. Tabelle 1 zeigt die Übersicht der verfügbaren Daten zu patientenrelevanten Endpunkten aus den eingeschlossenen Studien.

Aus der Studie **Brown 2011** [24] konnten Daten zur Mortalität und Morbidität (Infektionen) verwendet werden. Die Studie **Myers 2002** [25] berichtet für die Vergleichsgruppe nur Endpunkte für die Überlebenden nach der Transplantation. Daher waren für die Nutzenbewertung aus dieser Studie nur die Daten zum Endpunkt Mortalität verwertbar. Aus der Studie **Cuvelier 2016** [26] waren für die Nutzenbewertung Daten zur Morbidität heranzuziehen (Infektionen zum Zeitpunkt der Transplantation).

Daten zu den Endpunkten Krankenhausaufenthalte, Entwicklungsstörungen, unerwünschte Ereignisse und gesundheitsbezogene Lebensqualität konnten, sofern berichtet, aus keiner Studie verwertet werden.

Tabelle 1: Matrix der Endpunkte

Studie	Endpunkte							
	Mortalität	Infektionen	Graft-versus-Host Reaktion	Erneute Transplantation	Krankenhausaufenthalte	Entwicklungsstörungen	Unerwünschte Ereignisse	Gesundheitsbezogene Lebensqualität sowie psychosoziale Aspekte
Brown 2011	•	•						
Myers 2002	•							
Cuvelier 2016		•	(•) ^a	(•) ^a				
a: Da nur insgesamt 8 Kinder in dieser Studie untersucht wurden, kann sie in Bezug auf Endpunkte, zu denen keine weiteren Studien Daten berichtet haben, keine belastbare Datengrundlage für die Nutzenbewertung beitragen. Auf die Darstellung wird daher verzichtet.								

Aus 5 Studien zur diagnostischen Güte (**Kwan 2015** [27], **Kwan 2013** [23], **Chien 2015** [30], **Kanegae 2016** [31] und **de Felipe 2016** [32]) wurden Daten für die Nutzenbewertung herangezogen. Aufgrund des durchgängig gewählten VOPT-Designs konnte ausschließlich der positive prädiktive Wert (PPV) als Maß der diagnostischen Güte bewertet werden.

4.4 Bewertung des Verzerrungspotenzials auf Studien- und Endpunktebene

Das Verzerrungspotenzial auf Studienebene wurde für die 3 Studien zum **Vergleich früherer versus späteren Therapiebeginn** als hoch eingestuft, dies unter anderem deshalb, weil es sich nicht um randomisierte Studien handelte. Einer der weiteren Gründe lag in der

unzureichenden Beschreibung des zeitlichen Versatzes von Patienten der Interventions- und Kontrollgruppe.

Das hohe Verzerrungspotenzial auf Studienebene schlug sich direkt auf das Verzerrungspotenzial auf Endpunktebene (Mortalität und Morbidität) nieder.

Auch die **Studien zur diagnostischen Güte**, deren Ergebnisse für die Nutzenbewertung herangezogen werden konnten, wurden auf Studienebene jeweils mit einem hohen Verzerrungspotenzial bewertet. Dies lag zum einen daran, dass der in den Studien zur Überprüfung positiver Testergebnisse angewandte Referenztest, die Durchflusszytometrie, einen Beurteilungsspielraum lässt. Damit blieb unklar, inwieweit sich die Kenntnis des Screeningtestergebnisses auf die Interpretation des Referenztestes auswirkte. Hinzu kam übereinstimmend die jeweils unzureichende Beschreibung des Patientenflusses.

4.5 Ergebnisse zu patientenrelevanten Endpunkten

4.5.1 Ergebnisse zum Endpunkt Mortalität

Es liegen Daten aus 2 Studien vor.

Die Ergebnisse zur Mortalität aus der Studie **Brown 2011** [24] zeigen bei Neugeborenen mit positiver Familienhistorie einen dramatischen Effekt zugunsten des früheren Therapiebeginns im Sinne eines früheren Behandlungsbeginns mit Infektionsprophylaxe (nicht: Transplantationszeitpunkt; der ist in dieser Studie nicht berichtet) im Vergleich zu einem späteren Therapiebeginn. Berichtet werden insgesamt 6 Todesfälle bei 60 eingeschlossenen Patienten (10 %) in der Interventionsgruppe versus 29 Todesfälle bei 48 eingeschlossenen Patienten (60 %) in der Vergleichsgruppe. Der Gruppenunterschied zeigt mit einem Odds Ratio von 0,07 (95 %-KI [0,03; 0,20]) einen p-Wert von < 0,001. Es lässt sich ein Anhaltspunkt für einen Nutzen des früheren Behandlungsbeginns im Sinne einer Infektionsprophylaxe, gefolgt von einer Stammzelltransplantation, ableiten.

In dieser Studie sind die Ergebnisse danach aufgeschlüsselt, wann die Patienten verstorben sind (vor oder nach der Transplantation). Vor der Transplantation verstarb 1 Patient (1,8 %) in der Interventionsgruppe im Vergleich zu 17 Patienten (35 %) in der Vergleichsgruppe, womit der Gruppenunterschied hier bei einem Odds Ratio von 0,03 (95 %-KI [0,00; 0,24]) und wiederum einem p-Wert von < 0,001 liegt. Für den Zeitpunkt nach der Transplantation errechnet sich ein Odds Ratio von 0,27 (95 %-KI [0,09; 0,84]) und ein p-Wert von 0,019.

Die Ergebnisse aus der Studie **Myers 2002** [25] zeigen zwar keinen dramatischen Effekt, gleichwohl zeigt sich – ebenfalls bei Neugeborenen mit positiver Familienhistorie – ein deutlicher Vorteil einer früheren im Vergleich zu einer späteren Transplantation (< / > 3,5 Lebensmonate). Von 21 Patienten der Interventionsgruppe verstarb 1 Patient (5 %); in der Vergleichsgruppe von 96 Patienten verstarben 25 Patienten (26 %). In dieser Studie zeigt der Gruppenunterschied einen p-Wert von 0,080 und ein Odds Ratio von 0,14 (95 %-KI [0,02; 1,11]). Diese Ergebnisse stützen den bei Brown 2011 gezeigten Effekt eines früheren

Behandlungsbegins, auch wenn bei Myers 2002 der Zeitpunkt der Transplantation und nicht derjenige der Infektionsprophylaxe Gegenstand der Untersuchung war.

Der gepoolte Gesamteffekt beider Studien zeigt einen p-Wert von $< 0,001$ (Odds Ratio 0,08, 95 %-KI [0,03; 0,21]) und ist damit statistisch signifikant zugunsten eines früheren Therapiebeginns. Diese Gesamtbetrachtung lässt außer Acht, dass in beiden Studien der Beginn unterschiedlicher Therapien (Zeitpunkt der Infektionsprophylaxe einerseits und Zeitpunkt der Stammzelltransplantation andererseits) untersucht wurde.

Die Ergebnissicherheit der Studien war aufgrund des Studiendesigns und des hohen Verzerrungspotenzials als gering einzustufen. Es handelte sich um retrospektiv aus der Praxis erhobene Daten mit unzureichender Beschreibung des zeitlichen Versatzes bezüglich des Patienteneinschlusses in die beiden Gruppen, der Dauer der Nachbeobachtung, des Patientenflusses und der Vollständigkeit der Daten. Insgesamt war für den früheren Therapiebeginn ein Anhaltspunkt für einen Nutzen abzuleiten, bezogen auf den früheren Beginn einer in eine kurative Therapie mündenden infektionsprophylaktischen Behandlung.

4.5.2 Ergebnisse zum Endpunkt Morbidität (Infektionen)

In den Bericht sind Daten aus 2 Studien eingeflossen. Daten aus der Studie **Brown 2011** [24] zur Morbidität liegen nur unvollständig vor (Infektionen; der konkrete Zeitpunkt der Infektionen ist nicht angegeben). In der Interventionsgruppe wurden die Daten für 3 von 60 Patienten als „nicht verfügbar“ angegeben, in der Vergleichsgruppe für 19 von 48 Patienten. Gleichwohl konnte eine eigene (konservative) Sensitivitätsanalyse gerechnet werden, in der die Ereignisse auf die vollständigen Gruppen und nicht nur auf die tatsächlich nachbeobachteten Patienten bezogen wurden. Damit waren aus der Interventionsgruppe mit 60 Patienten 10 Patienten mit Infektionen der Vergleichsgruppe mit 48 Patienten und 26 Ereignissen gegenüberzustellen. Dieser Vergleich zeigt einen statistisch signifikanten Gruppenunterschied mit einem p-Wert $< 0,001$ bei einem Odds Ratio von 0,17 (95 %-KI [0,07; 0,41]).

Aus der Studie **Cuvelier 2016** [26] wurden Daten zu Infektionen zum Zeitpunkt der Transplantation herangezogen. Hier bestand die Interventionsgruppe aus 2 Kindern jeweils ohne Infektion zum Zeitpunkt der Stammzelltransplantation und die Vergleichsgruppe aus 6 Kindern, von denen 5 zum Zeitpunkt der Transplantation einen Infekt hatten. Der Vergleich zeigt einen statistisch nicht signifikanten Gruppenunterschied mit einem p-Wert von 0,055 bei einem Odds Ratio von 0,05 (95 %-KI [0; 1,87]).

Bezogen auf den Endpunkt Morbidität ist keine Nutzensaussage möglich.

4.6 Ergebnisse zur diagnostischen Güte

Aus 5 von 8 eingeschlossenen Studien zur diagnostischen Güte wurden die Daten zur Berechnung des PPV herangezogen. Angaben zur Berechnung der Sensitivität und Spezifität fehlen.

Aus den Angaben in **Kwan 2015** wurde für den Indextest ein PPV von 100 % berechnet (95 %-KI [39,8; 100], bei 4 richtig-positiven von 4 positiven Testergebnissen). Wesentlich niedriger liegt der PPV, der aus den anderen verwertbaren Studien berechnet wurde. Aus **Kwan 2013** wurde ein PPV von 7,5 % berechnet (95 %-KI [3,9; 12,7], bei 12 richtig-positiven von 161 positiven Testergebnissen). Der PPV aus den Daten von **Chien 2015** liegt bei 8,3 % (95 %-KI [1,0; 27,0], bei 2 richtig-positiven von 24 positiven Testergebnissen). Einen PPV von 0 zeigen **Kanegae 2016** (95 %-KI [0; 70,8]) und **de Felipe 2016** (95 %-KI [0; 60,2]) mit jeweils 0 richtig-positiven Testergebnissen von 3 bzw. 4 nachbeobachteten positiven Testergebnissen.

Die Ergebnisse der ausgewerteten Studien sind heterogen. Zu beachten ist, dass die Navajo-Population aus der Studie **Kwan 2015** eine 30-fach höhere Prävalenz für SCID aufweist als „übliche“ Populationen [27], und auch **Chien 2015** berichtet von einer vergleichsweise hohen Prävalenz für SCID in Taiwan. Die Ergebnisse der mit Abstand größten Studie **Kwan 2013** entsprechen am ehesten dem, was bei einer Anwendung des Tests in Deutschland erwartet werden kann.

Sensitivitäts- oder Subgruppenanalysen wurden nicht durchgeführt.

4.7 Studien unklarer Relevanz

Über die Studienregisterrecherche wurde 1 in Frankreich laufende multizentrische nicht randomisierte Vergleichsstudie identifiziert [34]. Diese untersucht zum einen die Testgüte der TREC-Messung aus Filterkartenblut, wobei über Transplantationszentren und Sterberegister auch die Sensitivität des Tests festgestellt werden soll. Als Zweites wird der Effekt des SCID-Screenings evaluiert, indem die Behandlungsergebnisse von Kindern verglichen werden, bei denen mit oder ohne ein Screening ein SCID festgestellt wird. Die Datenerhebung soll Mitte 2018 abgeschlossen werden.

4.8 Landkarte der Beleglage

Auf die Darstellung der Landkarte der Beleglage wird aufgrund der übersichtlichen Evidenzlage verzichtet.

Die Ergebnisse zum Endpunkt Mortalität zeigen einen Anhaltspunkt für einen Nutzen. Auf Basis der Ergebnisse zum Endpunkt Morbidität ist eine Nutzensaussage nicht möglich, gleichwohl stützen sie dieses Ergebnis.

Die Ergebnisse der Studien zur Testgüte zeigen, dass ein TREC-basierter Test auf SCID geeignet ist, Neugeborene mit SCID zu identifizieren, aber auch falsch-positive Befunde erbringt. Daten zur Anzahl falsch-negativer Ergebnisse liegen nicht vor.

5 Einordnung des Arbeitsergebnisses

Der schwere kombinierte Immundefekt ist eine seltene und schwere Erkrankung. Die Methodik zur Bewertung des Nutzens eines Screenings auf SCID wurde daher durch die Möglichkeit der Nutzenbewertung auch mittels Bewertung der einzelnen Bausteine der Screeningkette und durch den Einschluss von Studien niedriger Evidenzstufen weit gefasst.

Angesichts der Tatsache, dass seit 2008 sukzessive in 39 Bundesstaaten in den USA ein SCID-Screening in die Neugeborenenversorgung implementiert wurde (Stand Juni 2016 [35]) und parallel Register mit SCID-Fällen geführt werden [35,36], war zu erwarten, dass Ergebnisse des SCID-Screenings entweder mit Daten aus nicht screenenden Bundesstaaten oder auch historisch verglichen werden. Gleichwohl wurde nur 1 vergleichende Interventionsstudie der Screeningkette identifiziert; deren Daten waren nicht interpretierbar.

Zur Nutzenbewertung wurden daher die einzelnen Bausteine der Screeningkette bewertet (Linked-Evidence-Ansatz).

2 der eingeschlossenen Studien zum Vergleich zwischen früherem und späterem Behandlungsbeginn verglichen die Therapieergebnisse jeweils zwischen einem Neugeborenen mit positiver SCID-Familienanamnese und dem älteren Verwandten, bei dem kein SCID-Verdacht bestand und SCID daher erst nach Symptomeintritt diagnostiziert wurde. Bei einer weiteren Studie bestand die Interventionsgruppe ebenfalls aus aufgrund positiver Familienhistorie asymptomatisch identifizierten Kindern. Auch wenn das Testen auf SCID bei familiär bekanntem SCID-Risiko kein SCID-Screening darstellt, können die Studienergebnisse dennoch als übertragbar auf die Screeningsituation gelten, weil mehrfach in einer Familie auftretende SCID-Fälle sich in Bezug auf den Schweregrad, den Verlauf und die Therapie offenbar nicht von einzeln (oder erstmalig) auftretenden SCID-Fällen unterscheiden. Als Zweites muss darauf hingewiesen werden, dass 1 Studie einen Nutzen eines früheren Behandlungsbeginns (infektionsprophylaktische Therapie) anzeigt, dass dies aber nicht bedeutet, dass SCID-Patienten von einer möglichst frühen *kurativen* Therapie (in der Regel Knochenmark- oder Stammzelltransplantation) profitieren. Die Frage nach dem optimalen Transplantationszeitpunkt und der Notwendigkeit einer Chemotherapie bei Neugeborenen mit SCID wird derzeit kontrovers diskutiert [3,37,38].

Dass für die frühzeitige Infektionsprophylaxe sofort nach einer SCID-Diagnose zur Vermeidung schwerer oder gar tödlicher Infektionen lediglich ein Anhaltspunkt für einen Nutzen abgeleitet werden konnte, liegt an den Unklarheiten der ausschlaggebenden Studie **Brown 2011** [24]. Neben dem nicht randomisierten Design sind weitere methodische Probleme der Studie zu bedenken, nämlich die möglicherweise erhebliche zeitliche Nichtparallelität der Gruppen, die fehlenden Angaben zum Transplantationszeitpunkt und die unklare Dauer der Nachbeobachtung.

Die Datenlage aus den Studien zur diagnostischen Güte reicht nicht aus, um die Sensitivität und Spezifität zu berechnen; die Anzahl der falsch- und richtig-negativen Befunde ist unklar.

Die aus den eingeschlossenen Studien ermittelten PPV weisen eine maximale Spannweite auf, wobei für den deutschen Versorgungskontext davon auszugehen ist, dass in Anlehnung an die mit Abstand größte Studie Kwan 2013 mit dem Screening ein PPV im Bereich zwischen 4 % und 13 % zu erwarten ist.

Die eingeschlossenen Studien zur diagnostischen Güte beschränken sich mit 1 Ausnahme auf einen SCID-Test mittels TREC-Quantifizierung. Die Studie zum kombinierten Test auf TREC und KREC [32] zeigt mangels richtig-positiver Testergebnisse einen PPV von 0. Eine weitere Studie zu dem kombinierten Test, die durch einen der Autoren unaufgefordert überlassen wurde [39], wurde nicht eingeschlossen, weil dies einer systematischen, nicht selektiven Informationsbeschaffung widersprochen hätte [40]. Die Autoren berichten von 3 Neugeborenen mit SCID, die in einem kombinierten TREC- und KREC-Screening von knapp 60 000 Neugeborenen identifiziert worden waren. Sie berichten weiter von 61 Fällen, in denen sich der SCID-Verdacht nach positivem Testergebnis nicht bestätigt hatte. Die Ergebnisse dieser noch nicht veröffentlichten Studie liegen im Rahmen der eingeschlossenen Studien zur diagnostischen Güte. Ergebnisse weiterer Studien zum kombinierten TREC- und KREC-Test sind zu erwarten.

Eine Gegenüberstellung der gesundheitsbezogenen Konsequenzen von falsch-positiven, richtig-positiven, falsch-negativen sowie richtig-negativen Befunden war aufgrund der Datenlage nicht möglich. Falsch-positive Screeningbefunde, die grob geschätzt auf Basis der vorhanden Daten mit einer Wahrscheinlichkeit von etwa 1:10 zu erwarten sind, können (psychische) Belastungen der Eltern für den Zeitraum bis zur Abklärung durch Vollblutanalysen und eine genetische Analyse auslösen. Die Vor- und Nachteile einer früheren Behandlung von Kindern mit anderen Pathologien als SCID, die durch das SCID-Screening identifiziert wurden, können auf Basis der eingeschlossenen Studien nicht beurteilt werden. Schädliche Folgen falsch-negativer Testergebnisse sind dann denkbar, wenn Eltern oder medizinisches Personal bei anderweitigen Anzeichen für SCID diesem Verdacht wegen des unauffälligen Testergebnisses nicht mithilfe weiterer abklärender Diagnostik nachgehen. Bei falsch-negativem Untersuchungsergebnis wäre es mithin möglich, dass notwendige medizinische Interventionen verzögert werden. Allerdings sollte von dieser Konstellation in der Praxis nicht auszugehen sein. Weil ein negatives Testergebnis Neugeborene nicht von den übrigen Untersuchungen nach bisherigem diagnostischem Standard ausschließt, sollten Neugeborene mit falsch-negativem Ergebnis aus einem SCID-Screening keinen Schaden zu erwarten haben im Vergleich zu einer Diagnosestrategie ohne SCID-Screening.

Insgesamt ist im deutschen Versorgungskontext bei Zugrundelegung der vorhandenen Daten zur Prävalenz (siehe Kapitel 1) davon auszugehen, dass ein SCID-Screening aller Neugeborenen mittels TREC aus Filterkartenblut bei jährlich 700 000 Geburten bestenfalls alle etwa 20 bis 30 erkrankten Neugeborenen identifizieren könnte, von denen etwa ein Drittel (geschätzt mit sehr großer Unsicherheit) durch einen früheren Behandlungsbeginn vor dem Tod bewahrt werden könnte.

Für das Vereinigte Königreich wurde keine Empfehlung für die Einführung des Screenings ausgesprochen, und zwar aufgrund fehlender Informationen zur Epidemiologie, zur Testgüte und zum Umgang mit falsch-positiven Befunden sowie aufgrund der offenen Frage der klinischen und der Kosteneffektivität [4] – obwohl ein Vorteil einer früheren Therapie festgestellt wurde. Überlegungen laufen noch in verschiedenen europäischen Ländern; in Frankreich [34], Spanien [32,33], Norwegen, Schweden und den Niederlanden laufen derzeit Pilotprojekte oder sind angekündigt.

6 Fazit

Das Neugeborenencreening auf schwere kombinierte Immundefekte (SCID-Screening) in Kombination mit einer infektionsprophylaktischen Therapie, die in eine kurative Therapie mündet, zeigt einen Anhaltspunkt für einen Nutzen.

Für den Endpunkt Mortalität wird ein dramatischer Effekt zugunsten eines früheren versus einen späteren Beginn einer infektionsprophylaktischen Therapie beschrieben. Die Datenlage aus den Studien zur diagnostischen Güte reichte nicht aus, um die Sensitivität und Spezifität zu berechnen. Der positive prädiktive Wert (PPV) für ein SCID-Screening durch die Bestimmung von T-Cell Receptor Excision Circles (TREC) aus Filterkartenblut zeigt, dass der Test auf SCID grundsätzlich geeignet ist, Kinder mit SCID zu identifizieren. Unklar ist, wie viele Kinder nicht gefunden werden.

Diese Ergebnisse stützen sich auf 3 nicht randomisierte vergleichende Interventionsstudien zum Therapiebeginn sowie 5 Studien zur diagnostischen Güte, die mittels Linked-Evidence-Ansatz zusammengeführt wurden. Sämtliche Studien weisen ein hohes Verzerrungspotenzial auf.

Eine in Frankreich laufende vergleichende Interventionsstudie der Screeningkette, in der auch Daten zur Sensitivität und Spezifität erhoben werden sollen, lässt Ergebnisse zur Beantwortung der Nutzenfrage des SCID-Screenings erwarten; Daten werden bis Juni 2018 erhoben.

Details des Berichts

A1 Projektverlauf

A1.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 30.07.2015 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung eines Screenings auf schwere kombinierte Immundefekte (SCID-Screening) bei Neugeborenen beauftragt.

Der Beschluss zur Auftragserteilung des G-BA nennt als Gegenstand der Nutzenbewertung ein Screening auf „schwere angeborene Immundefekte (SCID-Screening) bei Neugeborenen“. Aus der Konkretisierung des Auftrags des G-BA vom 30.07.2015 geht hervor, dass sich die Nutzenbewertung auf ein Screening auf schwere kombinierte Immundefekte beziehen soll. Da das im Beschluss zur Auftragserteilung verwendete Akronym SCID (Severe combined Immunodeficiency) ebenfalls kombinierte Immundefekte beschreibt, bezieht sich auch die Nutzenbewertung auf angeborene kombinierte Immundefekte.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 24.09.2015 Betroffene zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen konsultiert.

Der vorläufige Berichtsplan in der Version 1.0 vom 23.11.2015 wurde am 10.12.2015 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 13.01.2016 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 18.02.2016) publiziert.

Die vorläufige Bewertung, der Vorbericht in der Version 1.0 vom 21.07.2016, wurde am 29.07.2016 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 26.08.2016 konnten schriftliche Stellungnahmen eingereicht werden. Die in den Stellungnahmen vorgebrachten Argumente werden im Kapitel „Kommentare“ des vorliegenden Abschlussberichts gewürdigt.

Der vorliegende Abschlussbericht beinhaltet die Änderungen, die sich aus der Anhörung ergeben haben.

Im Anschluss an die Anhörung erstellte das IQWiG den vorliegenden Abschlussbericht, der 8 Wochen nach Übermittlung an den G-BA auf der Website des IQWiG veröffentlicht wird. Die zum Vorbericht eingegangenen Stellungnahmen werden in einem gesonderten Dokument

„Dokumentation und Würdigung der Anhörung zum Vorbericht“ zeitgleich mit dem Abschlussbericht im Internet bereitgestellt.

A1.2 Dokumentation der Änderungen im Projektverlauf

Berichtsplan im Vergleich zum vorläufigen Berichtsplan

- Konferenzabstracts wurden im Rahmen der bibliografischen Recherche ausgeschlossen.

Im Abschnitt „Bibliografische Recherche“ (A2.1.5.1) wurde ergänzt, dass Konferenzabstracts im Rahmen der bibliografischen Recherche ausgeschlossen werden.

Vorbericht im Vergleich zum Berichtsplan

Etwaige methodische Spezifizierungen und Änderungen werden in Abschnitt A2.2 beschrieben.

Darüber hinaus ergaben sich lediglich redaktionelle Änderungen.

Abschlussbericht im Vergleich zum Vorbericht

Etwaige methodische Spezifizierung und Änderungen werden detailliert in Abschnitt A2.2 beschrieben.

Neben redaktionellen Änderungen ergaben sich darüber hinaus folgende Änderungen:

- In Kapitel 1 wurde aufgrund einer Stellungnahme der Verweis auf ein zwischenzeitlich zugelassenes Gentherapeutikum zur Behandlung von ADA-SCID aufgenommen.
- In Tabelle 17 wurde der Eintrag zum Verzerrungspotenzial („keine sonstigen Aspekte“) korrigiert. Eine inhaltliche Änderung der Bewertung ist damit nicht verbunden.
- Im Rahmen der Nachrecherche wurden 2 neu einzuschließende Studien identifiziert und im Abschlussbericht berücksichtigt. Entsprechende Ergänzungen finden sich in den Kapiteln 4, 5, 6 und A3. Die Ergebnisse der neu eingeschlossenen Studien hatten keine inhaltlichen Auswirkungen auf das Fazit.
- Nach Abschluss der systematischen Recherche ging die Antwort auf eine zum Vorbericht versandte Autorenanfrage ein (de Felipe 2016 [32]). Die Ergebnisse dieser Studie wurden im Abschlussbericht ebenfalls berücksichtigt. Auch diese Ergebnisse hatten keine inhaltlichen Auswirkungen auf das Fazit.
- In Kapitel 5 wird auf ein eingereichtes Studienmanuskript hingewiesen, das durch einen der Autoren unaufgefordert überlassen wurde [39]. Die Studie wurde nicht eingeschlossen, weil dies einer systematischen, nicht selektiven Informationsbeschaffung widersprochen hätte.

A2 Details der Methoden

A2.1 Methodik gemäß Berichtsplan

Nutzenbewertung anhand von vergleichenden Interventionsstudien der Screeningkette

Der Nutzen von Screeningmaßnahmen lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien der gesamten Screeningkette mit einer (idealerweise randomisierten) Zuteilung von Personen zu einer Strategie mit beziehungsweise ohne Anwendung der Screeningmaßnahme und der Betrachtung patientenrelevanter Endpunkte bewerten [40]. In einer solchen Studie erhalten die Neugeborenen in der Kontrollgruppe kein SCID-Screening. Die der Interventionsgruppe zugeteilten Neugeborenen erhalten ein SCID-Screening und werden entsprechend dem Untersuchungsergebnis des SCID-Screenings gegebenenfalls diagnostischer Abklärung und Therapie zugewiesen.

Liegen weder randomisiert kontrollierte Studien (RCTs) noch nicht randomisierte Interventionsstudien vor, können im Rahmen dieser Fragestellung vergleichende Studien mit niedrigerem Evidenzniveau hinzugezogen werden. Auf Basis dieser Studien sind Nutzensaussagen nur möglich, wenn dramatische Effekte vorliegen.

Nutzenbewertung anhand von vergleichenden Studien zum Therapiebeginn und Studien zur Bewertung der diagnostischen Güte

Liegen vergleichende Interventionsstudien der Screeningkette für die Nutzenbewertung nicht oder in nicht ausreichender Quantität und Qualität vor, kann eine Bewertung der einzelnen Bausteine der Screeningkette erfolgen. Für die Nutzenbewertung werden gesundheitsbezogene Konsequenzen für falsch-positive, richtig-positive, falsch-negative sowie richtig-negative Befunde gegenübergestellt, der Nutzen eines frühen gegenüber einem späten Therapiebeginn erfasst und die diagnostische Güte untersucht. Der Nutzen des Screenings kann dann dadurch abgeleitet werden, dass ein früherer gegenüber einem späteren Therapiebeginn Nutzen zeigt und gleichzeitig der Screeningtest entsprechend der Gegenüberstellung eine hinreichende diagnostische Güte aufweist. Dazu werden vergleichende Interventionsstudien zum Therapiebeginn und Studien zur diagnostischen Güte herangezogen. Liegen weder RCTs noch nicht randomisierte Interventionsstudien zur Vorverlegung des Therapiebeginns vor, können vergleichende Studien mit niedrigerem Evidenzniveau hinzugezogen werden. Auf Basis dieser sind Nutzensaussagen nur möglich, wenn dramatische Effekte vorliegen.

A2.1.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette in die Untersuchung

A2.1.1.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

A2.1.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention ist das SCID-Screening bei Neugeborenen. Dafür werden die im Rahmen der Neugeborenenuntersuchung eingesetzten Filterpapierkarten verwendet. Der Zeitpunkt der Probenentnahme soll auf den in den Kinder-Richtlinien genannten Zeitrahmen übertragbar sein. Die laboranalytische Methodik und der Grenzwert zur Unterscheidung positiver und negativer Ergebnisse müssen prospektiv festgelegt worden sein. Beide müssen an einer unabhängigen Stichprobe entwickelt worden sein. Als Vergleichsintervention gilt kein Screening. Die Therapie besteht in einer allogenen Knochenmark- oder Stammzelltransplantation zur Rekonstitution einer intakten Immunfunktion (siehe Kapitel 1).

A2.1.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtüberleben, krankheitsspezifisches Überleben),
- Morbidität (z. B. Infektionen oder Organschäden, Transplantationen),
- Krankenhausaufenthalte (insbesondere auch solche in Umkehrisolierung),
- Entwicklungsstörungen (z. B. Störungen der kognitiven, psychosozialen, emotionalen, grob- und feinmotorischen Entwicklung),
- unerwünschte Ereignisse,
- gesundheitsbezogene Lebensqualität des Kindes (gemessen z. B. durch Proxy Rating).

Subjektive Endpunkte (zum Beispiel gesundheitsbezogene Lebensqualität) werden nur dann berücksichtigt, wenn sie mit validen Messinstrumenten (zum Beispiel validierten Skalen) erfasst wurden.

A2.1.1.4 Studientypen

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 oder Zusatznutzens einer medizinischen Intervention.

Aufgrund der Seltenheit von SCID scheinen RCTs und nicht randomisierte Interventionsstudien für die unter A2.1.1.2 genannte Intervention kaum erwartbar. Gleichzeitig erscheint es denkbar, dass die zu prüfende Intervention einen dramatischen Effekt aufweist, der sich auch in Studien mit niedrigerem Evidenzniveau nicht allein durch Verzerrung erklären lässt. Wenn die auf RCTs und nicht randomisierten Interventionsstudien basierende Datenlage nicht reicht, um den patientenrelevanten Nutzen und Schaden des SCID-Screenings bei Neugeborenen mit ausreichender Ergebnissicherheit schätzen zu können, werden zu dieser Fragestellung daher auch vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) als relevante wissenschaftliche Literatur in die Nutzenbewertung

einfließen. Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber gegebenenfalls stärken. Auf Basis solcher Studien sind Nutzensaussagen nur möglich, wenn dramatische Effekte vorliegen.

A2.1.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.1.6 Tabellarische Darstellung der Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

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

Einschlusskriterien	
INS1	Neugeborene (siehe auch Abschnitt A2.1.1.1)
INS2	Prüfintervention: Neugeborenencreening auf SCID unter Verwendung von Filterpapierkarten (siehe auch Abschnitt A2.1.1.2)
INS3	Vergleichsintervention: kein Neugeborenencreening auf SCID (siehe auch Abschnitt A2.1.1.2)
INS4	patientenrelevante Endpunkte wie in Abschnitt A2.1.1.3 formuliert
INS5	Studientypen: RCT, nicht randomisierte Interventionsstudien, Kohortenstudien (auch retrospektiv oder mit historischem Vergleich)
INS6	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [41] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [42], TREND- [43] oder STROBE-Statements [44] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Conference of Harmonization; RCT: Randomized controlled Trial (randomisierte kontrollierte Studie); SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs.</p>	

A2.1.2 Kriterien für den Einschluss von vergleichenden Interventionsstudien zum Therapiebeginn in die Untersuchung

Studien, die einen früheren versus einen späteren Therapiebeginn vergleichen, werden im Rahmen des vorliegenden Berichts systematisch recherchiert und ausgewertet, wenn vergleichende Interventionsstudien zur Screeningkette nicht oder in nicht ausreichender Quantität und Qualität vorliegen.

A2.1.2.1 Population

In die Bewertung werden Studien mit Patienten mit SCID aufgenommen. Die Diagnosestellung bei Patienten mit früherem Therapiebeginn muss auf die Screeningsituation übertragbar sein.

A2.1.2.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet ein früherer Therapiebeginn. Als Vergleichsintervention gilt ein späterer Therapiebeginn. Bezüglich der Definitionen von „früherem Therapiebeginn“ und „späterem Therapiebeginn“ bestehen keine Einschränkungen, jedoch dürfen die eingesetzten Therapiemethoden zwischen Prüf- und Vergleichsintervention keine wesentlichen Unterschiede aufweisen. Die Therapie besteht in einer allogenen Knochenmark- oder Stammzelltransplantation zur Rekonstitution einer intakten Immunfunktion (siehe Kapitel 1).

A2.1.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden die unter Abschnitt A2.1.1.3 genannten patientenrelevanten Endpunkte betrachtet.

A2.1.2.4 Studientypen

Aufgrund der Seltenheit von SCID scheinen RCTs und nicht randomisierte Interventionsstudien für die unter Abschnitt A2.1.2.2 genannte Intervention und alle unter Abschnitt A2.1.2.3 genannten Endpunkte kaum erwartbar. Gleichzeitig erscheint es denkbar, dass die zu prüfende Intervention mit früherem Therapiebeginn einen dramatischen Effekt gegenüber der Vergleichsintervention mit späterem Therapiebeginn aufweist, der sich auch in vergleichenden Studien mit niedrigerem Evidenzniveau nicht allein durch Verzerrung erklären lässt. Daher werden zu dieser Fragestellung auch vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen. Auf Basis dieser sind Nutzensaussagen nur möglich, wenn dramatische Effekte vorliegen.

A2.1.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.2.6 Tabellarische Darstellung der Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)

Einschlusskriterien	
INT1	Patienten mit SCID (siehe auch Abschnitt A2.1.2.1)
INT2	Prüfintervention: frühere Behandlung (siehe auch Abschnitt A2.1.2.2)
INT3	Vergleichsintervention: spätere Behandlung (siehe auch Abschnitt A2.1.2.2)
INT4	patientenrelevante Endpunkte wie in Abschnitt A2.1.2.3 formuliert
INT5	Studientypen: Kohortenstudien (auch retrospektiv oder mit historischem Vergleich)
INT6	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [41] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [42], TREND- [43] oder STROBE-Statements [44] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Conference of Harmonization; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

A2.1.3 Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Untersuchung

A2.1.3.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

A2.1.3.2 Indextest

Indextest ist die Testung auf SCID unter Verwendung von Filterpapierkarten. Der Zeitpunkt der Probenentnahme soll auf den in den Kinder-Richtlinien des G-BA genannten Zeitrahmen übertragbar sein. Die laboranalytische Methodik und der Grenzwert zur Unterscheidung positiver und negativer Ergebnisse müssen prospektiv festgelegt worden sein. Beide müssen an einer unabhängigen Stichprobe entwickelt worden sein.

A2.1.3.3 Referenztest

Referenztest sind die Lymphozytenphänotypisierung beziehungsweise (bei ADA- und PNP-Mangel) Enzymbestimmungen (siehe Kapitel 1) und / oder die genetische Analyse sowie bei unauffälligem Befund zusätzlich die Nachbeobachtung.

A2.1.3.4 Zielgrößen

Eingeschlossen werden Studien, aus denen personenbezogene Daten zur Berechnung der diagnostischen Güte im Hinblick auf die Entdeckung von SCID ableitbar sind.

A2.1.3.5 Studientypen

Um die diagnostische Güte des Indextests zur Erkennung von SCID bei Neugeborenen möglichst unverzerrt bestimmen zu können, soll eine Gruppe von Neugeborenen, die zu einem bestimmten Zeitpunkt prospektiv rekrutiert und auf SCID gescreent wurde, zeitnah (Querschnittsdesign) mit dem Referenztest (nach-)untersucht beziehungsweise bei unauffälligem Befund nachbeobachtet werden. Dabei sind ein konsekutiver Einschluss der Neugeborenen und die Dokumentation der fehlenden Werte notwendig.

Ist die Datenlage aus solchen Studien unzureichend, werden in die vorliegende Bewertung aufgrund der Seltenheit von SCID sowohl diagnostische retrospektive Kohorten- als auch diagnostische Fall-Kontrollstudien aufgenommen.

Ist die Datenlage aus Studien, die sowohl positive als auch negative Ergebnisse mit dem Referenztest direkt überprüfen (komplette Verifikation) unzureichend, können Studien im VOPT-Design (VOPT: Verification of only positive Testers) herangezogen werden. Dabei werden alle positiven Ergebnisse im Indextest mit dem Referenztest untersucht [45]. Eine Bewertung der testnegativen Fälle und damit eine Bestimmung der Sensitivität oder Spezifität des Tests sind mit solchen Studien nicht möglich.

A2.1.3.6 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.3.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)

Einschlusskriterien	
D1	Neugeborene (siehe auch Abschnitt A2.1.3.1)
D2	Indextests: Testung auf SCID bei Neugeborenen unter Verwendung von Filterpapierkarten (siehe auch Abschnitt A2.1.3.2)
D3	Referenztest: Lymphozytenphänotypisierung beziehungsweise Enzymbestimmungen, genetische Analyse, Nachbeobachtung (siehe auch Abschnitt A2.1.3.3)
D4	Zielgrößen: personenbezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt A2.1.3.4)
D5	diagnostische Kohorten- und Fall-Kontrollstudien (siehe auch Abschnitt A2.1.3.5)
D6	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des STARD- [46] oder STROBE-Statements [44] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology</p>	

A2.1.4 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien INS1, INT1 und D1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium INS1, INT1 beziehungsweise D1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

Ebenfalls eingeschlossen werden Studien, die zu mindestens 80 % die Einschlusskriterien INS2, INT2 und D2 erfüllen (Prüfintervention, bezogen auf die Interventionsgruppe der Studie beziehungsweise Indextest bei Diagnosestudien) und zu mindestens 80 % das Einschlusskriterium INS3 und INT3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie).

A2.1.5 Informationsbeschaffung

A2.1.5.1 Bibliografische Recherche

Die systematische Recherche nach relevanten Studien wird in folgenden bibliografischen Datenbanken durchgeführt:

- Suche nach Primärstudien in den Datenbanken MEDLINE, Embase, Cochrane Central Register of Controlled Trials,
- Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und Embase parallel zur Suche nach relevanter Primärliteratur sowie Suche in den Datenbanken Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects und Health Technology Assessment Database.

Konferenzabstracts werden im Rahmen der bibliografischen Recherche ausgeschlossen.

A2.1.5.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien beziehungsweise Informationen zu relevanten Studien

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien beziehungsweise Informationen zu relevanten Studien zu ermitteln, werden weitere Quellen berücksichtigt. Die Rechercheergebnisse werden anschließend auf weitere relevante Studien und Studienunterlagen untersucht (siehe Abschnitt A2.1.5.3).

A2.1.5.2.1 Systematische Übersichten

Relevante systematische Übersichten werden hinsichtlich weiterer relevanter Publikationen beziehungsweise Studien gesichtet.

A2.1.5.2.2 Öffentlich zugängliche Studienregister

Die folgenden öffentlich zugänglichen Studienregister werden durchsucht:

- U.S. National Institutes of Health. ClinicalTrials.gov, World Health Organization.
- International Clinical Trials Registry Platform Search Portal.

A2.1.5.2.3 Durch den G-BA übermittelte Unterlagen

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

A2.1.5.2.4 Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen

Es werden Anfragen an Autoren gestellt, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten lassen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen sind.

A2.1.5.2.5 Informationen aus Anhörungen

Im Anschluss an die Veröffentlichungen des vorläufigen Berichtsplans und des Vorberichts erfolgt eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen kann. Relevante Informationen aus diesen Anhörungen werden im Rahmen der Nutzenbewertung berücksichtigt.

A2.1.5.3 Selektion relevanter Studien

Selektion relevanter Publikationen aus den Ergebnissen der bibliografischen Recherche

Die durch die Suche in bibliografischen Datenbanken identifizierten und zu screenenden Treffer werden in einem ersten Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 2, Tabelle 3 und Tabelle 4) bewertet. Als potenziell relevant erachtete Publikationen werden in einem zweiten Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Reviewer unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden Reviewern aufgelöst.

Selektion relevanter Studien aus weiteren Suchquellen

Informationen aus den folgenden Suchquellen werden von 2 Reviewern unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

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

Informationen aus der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht werden von einem Reviewer auf Studien gesichtet, der diese dann in Bezug auf ihre Relevanz bewertet; ein zweiter Reviewer überprüft den gesamten Prozess inklusive der Bewertungen.

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

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

A2.1.6 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 werden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft. Auf Grundlage der Ergebnissicherheit wird für jedes Ergebnis endpunktspezifisch eine zugehörige Aussagesicherheit abgeleitet.

A2.1.6.1 Bewertung von vergleichenden Interventionsstudien

Datenextraktion

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Besonderes Augenmerk wird auf die Abbildung unterschiedlicher Screeningstrategien sowie die Auswirkung des SCID-

Screenings auf andere als schwere kombinierte Immundefekte gerichtet, sofern sich den eingeschlossenen Studien Informationen hierzu entnehmen lassen.

Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wird für jede in die Nutzenbewertung eingeschlossene Studie bewertet, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu werden 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

Für randomisierte Studien wird anhand dieser Aspekte das Verzerrungspotenzial zusammenfassend 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 wird zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter (A) aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft. Falls 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.

Für nicht randomisierte vergleichende Studien wird in der Regel keine zusammenfassende Bewertung der Verzerrungsaspekte durchgeführt, da die Ergebnisse dieser Studien aufgrund der fehlenden Randomisierung generell ein hohes Verzerrungspotenzial besitzen.

A2.1.6.2 Bewertung von Studien zur diagnostischen Güte

Datenextraktion

Aus den Unterlagen zu den Studien, die die unter A2.1.3 beschriebenen Einschlusskriterien erfüllen, werden Informationen in standardisierte Tabellen extrahiert. Besonderes Augenmerk wird auf die Abbildung unterschiedlicher Screeningstrategien sowie die Auswirkung des SCID-Screenings auf andere als schwere kombinierte Immundefekte gerichtet, sofern sich den eingeschlossenen Studien Informationen hierzu entnehmen lassen.

Bewertung des Verzerrungspotenzials der Ergebnisse

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der Primärstudien zur diagnostischen Güte erfolgt auf Basis des QUADAS-2-Instruments [47]. Das Verzerrungspotenzial von Primärstudien zur diagnostischen Güte wird als „niedrig“ 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.

A2.1.7 Informationssynthese und -analyse

Die Informationen werden einer Informationssynthese und -analyse unterzogen. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt darüber hinaus in jedem Fall.

A2.1.7.1 Gegenüberstellung der Ergebnisse der Einzelstudien

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

In bestimmten Fällen werden einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt beziehungsweise nicht in die Nutzenbewertung einbezogen. 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, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist. In der Literatur werden zum Teil bereits Auswertungen, in denen 20 % der Patienten nicht berücksichtigt werden, als nicht mehr aussagekräftig betrachtet [48].

Ausnahmen von dieser Regel werden zum Beispiel dann gemacht, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [49].

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patienten zwischen den Gruppen größer als 15 Prozentpunkte ist.

A2.1.7.2 Meta-Analysen

A2.1.7.2.1 Meta-Analysen für vergleichende Interventionsstudien

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar sind, werden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung werden primär die Ergebnisse aus ITT-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Meta-Analysen erfolgen in der Regel auf Basis von Modellen mit zufälligen Effekten [50]. 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 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 [51].

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 [52]. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ($p \geq 0,2$ für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität wird stattdessen das Prädiktionsintervall dargestellt und die Ergebnisse werden nur in begründeten Ausnahmefällen gepoolt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt A2.1.7.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt A2.1.7.5).

A2.1.7.2.2 Meta-Analysen für Studien zur diagnostischen Güte

Die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle [53] aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen erfüllt sind, für die Testgütekriterien eine Meta-Analyse anhand der Sensitivität und Spezifität in einer bivariaten Meta-Analyse durchgeführt [54]. Die

Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [55,56]. 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. 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 [57].

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. Sollten Sensitivität und Spezifität nicht berechenbar sein, zum Beispiel weil nur Studien im VOPT-Design eingeschlossen wurden, wird der PPV dargestellt und metaanalytisch zusammengefasst.

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen untersucht.

A2.1.7.3 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (Zusatz-)Nutzens und (höheren) Schadens in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (Zusatz-)Nutzen oder (höheren) Schaden vor“ getroffen.

Die Aussagesicherheit richtet sich nach der Anzahl verfügbarer Studien, der qualitativen und quantitativen Sicherheit ihrer Ergebnisse sowie der Homogenität der Ergebnisse bei mehreren Studien. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie zu differenzieren. Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit. Ergebnisse nicht randomisierter vergleichender Studien haben eine geringe qualitative Ergebnissicherheit. Die regelhaft abzuleitende Aussagesicherheit ist Tabelle 5 zu entnehmen. Der Nutzen des Screenings kann durch die Gegenüberstellung der gesundheitsbezogenen Konsequenzen und ihrer Wahrscheinlichkeit zusammen mit einer Aussage zum Nutzen eines früheren Therapiebeginns abgeleitet werden. Die Aussagesicher-

heit bezüglich des Nutzens des Screenings berücksichtigt dann sowohl die Aussagesicherheit bezüglich des Nutzens eines früheren Therapiebeginns als auch das Verzerrungspotenzial bezüglich der diagnostischen Güte.

Aussagen zum Nutzen auf Basis von Studien mit niedrigerer Evidenzstufe sind nur in Verbindung mit einem dramatischen Effekt möglich. Allein auf Basis der diagnostischen Güte wird keine Nutzensaussage abgeleitet.

Tabelle 5: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			homogen	heterogen		
			Meta- Analyse statistisch signifikant	gleichgerichtete Effekte ^a		
				deutlich	mäßig	nein
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–
a: Gleichgerichtete Effekte liegen vor, wenn trotz Heterogenität eine deutliche oder mäßige Richtung der Effekte erkennbar ist.						

A2.1.7.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sind Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel der Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder der Wahl des Effektmaßes. Derartige Sensitivitätsanalysen erfolgen unabhängig von gegebenenfalls weiteren Analysen, mit denen die Ergebnissicherheit eines beobachteten Effekts bewertet wird.

Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufteffekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (Zusatz-)Nutzen attestiert wird.

A2.1.7.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt A2.1.7.4 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 ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen, auch um die Übertragbarkeit der Ergebnisse besser beurteilen zu können:

- Geschlecht,
- Alter,
- Ethnie,
- Art der Immundefizienz (zum Beispiel T^B⁺ und T^B⁻).

Sollten sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

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

A2.2 Spezifizierungen und Änderungen der Methodik

Das Studienregister EU Clinical Trials Register wurde durchsucht. Darüber hinaus wurden keine Spezifizierungen oder Änderungen vorgenommen.

Spezifizierungen und / oder Änderungen der Methoden im Abschlussbericht im Vergleich zum Vorbericht

Es wurden keine Spezifizierungen und / oder Änderungen vorgenommen.

A3 Details der Ergebnisse

A3.1 Informationsbeschaffung

A3.1.1 Primäre Suchquellen

A3.1.1.1 Bibliografische Recherche

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien zum Studieneinschluss.

Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A7.1. Die letzte Suche fand am 15.08.2016 statt.

Die Zitate der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A6.3.

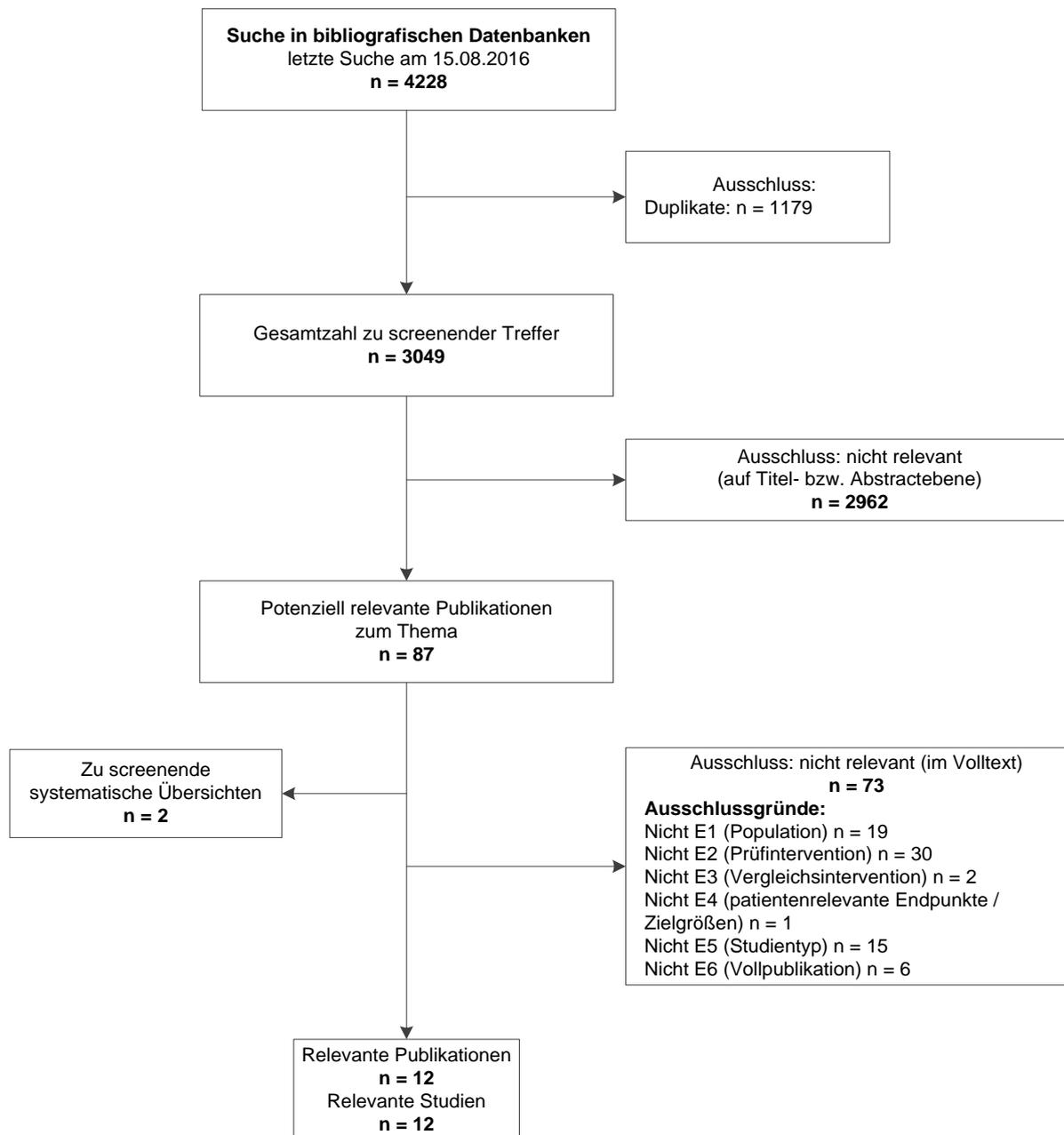


Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion

A3.1.1.2 Öffentlich zugängliche Studienregister

Durch die Suche in öffentlich zugänglichen Studienregistern wurden keine relevanten Studien bzw. Dokumente identifiziert.

Für die in Tabelle 6 dargestellten Studien konnte auf Basis der vorhandenen Informationen die Relevanz nicht abschließend geklärt werden. Es wurden keine Autorenanfragen gestellt.

Tabelle 6: In Studienregistern identifizierte Studien unklarer Relevanz^a

Studienregister ID	Studie	Studienregister	Status	Ergebnisbericht in Studienregister vorhanden
NCT02590328	Neonatal Screening of Severe Combined Immunodeficiencies	ClinicalTrials.gov [58]	not yet open for participant recruitment	nein
NCT00613561	Early Diagnosis and Stem Cell Transplantation for Severe Immunodeficiency Diseases (SIDS)	ClinicalTrials.gov [59]	recruiting participants	nein
NCT02244450	Generalized Neonatal Screening of Severe Combined Immunodeficiencies (DEPISTREC)	ClinicalTrials.gov [34]	recruiting participants	nein
NCT01186913	Natural History Study of SCID Disorders	ClinicalTrials.gov [60]	recruiting participants	nein
NCT01346150	Patients Treated for SCID (1968-2010)	ClinicalTrials.gov [61]	recruiting participants	nein
NCT00152100	Transplantation of Hematopoietic Cells in Children With Severe Combined Immunodeficiency Syndrome	ClinicalTrials.gov [62]	completed ^b	nein
<p>a: Eine Studie unklarer Relevanz ist eine Studie, für die keines der in Tabelle 2, Tabelle 3 und Tabelle 4 genannten Kriterien für den Studieneinschluss (ggf. mit Ausnahme des Vorliegens einer Vollpublikation) verletzt ist, aber auf Basis der vorliegenden Informationen nicht alle Kriterien eindeutig erfüllt sind.</p> <p>b: Der Eintrag bezieht sich auf eine Studie, die ggf. für den Vergleich eines früheren vs. einen späteren Therapiebeginn relevant sein könnte. Anhand des Studieneintrags ist nicht ersichtlich, ob die Studienpopulation auch solche Patienten enthält, die für die Nutzenbewertung relevant sind (Übertragbarkeit der Diagnosestellung auf die Screeningsituation).</p> <p>SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)</p>				

Die Suchstrategien für die Suche in Studienregistern finden sich in Abschnitt A7.2. Die letzte Suche in öffentlich zugänglichen Studienregistern fand am 16.08.2016 statt.

A3.1.2 Weitere Suchquellen

Über weitere Suchquellen identifizierte relevante Studien bzw. Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Suchquellen gefunden wurden.

A3.1.2.1 Systematische Übersichten

Im Rahmen der Informationsbeschaffung wurden 2 systematische Übersichten identifiziert – die entsprechenden Zitate finden sich in Abschnitt A6.2.

In diesen fanden sich keine relevanten Studien bzw. Dokumente, die nicht über andere Rechenschritte identifiziert werden konnten.

A3.1.2.2 Durch den G-BA übermittelte Dokumente

Im Rahmen der Auftragsbearbeitung wurden Dokumente vom G-BA an das IQWiG weitergeleitet. Diese wurden auf Duplikate zur bibliografischen Recherche überprüft. Alle verbleibenden Dokumente wurden bereits auf Titel- / Abstractebene ausgeschlossen.

Es fanden sich keine relevanten Studien bzw. Dokumente, die nicht über andere Rechenschritte identifiziert werden konnten.

A3.1.2.3 Anhörung

Es wurden keine relevanten Studien bzw. Dokumente genannt, die nicht über andere Rechenschritte identifiziert werden konnten.

A3.1.2.4 Autorenanfragen

Für die vorliegende Bewertung wurden Autorenanfragen versendet (Tabelle 7). Die Informationen aus den eingegangenen Antworten sind in die Studienbewertung eingeflossen.

Tabelle 7: Übersicht zu Autorenanfragen

Studie	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
De Felipe 2016	<ul style="list-style-type: none"> ▪ Klärung, ob die publizierten Daten auch Daten aus einer früher publizierten Pilotstudie enthalten ▪ falls ja, Bitte um Überlassung der um die Daten der Pilotstudie bereinigten Studienergebnisse 	ja	<ul style="list-style-type: none"> ▪ publizierte Daten enthalten auch Daten aus einer früher publizierten Pilotstudie ▪ Überlassung der um die Daten aus der Pilotstudie bereinigten Ergebnisse
Brown 2011	<ul style="list-style-type: none"> ▪ Geburtsjahr der eingeschlossenen Patienten ▪ patientenbezogene Dauer der Nachbeobachtung seit Diagnose, sofern nicht verfügbar als Median (Minimum / Maximum) ▪ Begründung der Diskrepanz berichteter Ergebnisse in der Probandenkohorte (Daten vollständig für den Endpunkt Mortalität, unvollständig für den Endpunkt Morbidität) 	ja	<ul style="list-style-type: none"> ▪ Autor lehnt die Beantwortung aus zeitlichen Gründen ab
Verbsky 2012	<ul style="list-style-type: none"> ▪ Zuordnung der Testergebnisse zum jeweils verwendeten Trennwert ▪ Erstellung eines Flussdiagramms geordnet nach jeweils verwendetem Trennwert inklusive Informationen zur Anzahl positiver Testergebnisse und durchgeführter sowie nicht durchgeführter Referenztests, Letzteres mit Angabe des Grundes für die Nichtdurchführung 	ja	<ul style="list-style-type: none"> ▪ Verweis auf die in der Publikation enthaltenen Daten
Chien 2015	<ul style="list-style-type: none"> ▪ prospektive Festlegung des Trennwertes ▪ Klärung von Differenzen in den Angaben zu als „inconclusive“ angegebenen Testergebnissen ▪ Klärung von Differenzen in den Angaben zur Anzahl getesteter Neugeborener und Anzahl von Testergebnissen 	ja	<ul style="list-style-type: none"> ▪ Trennwert wurde prospektiv festgelegt ▪ Erläuterung der Differenzen ▪ Erläuterung der Differenzen

A3.1.3 Resultierender Studienpool

Durch die verschiedenen Suchschritte konnten insgesamt 12 relevante Studien (12 Dokumente) identifiziert werden (siehe auch Tabelle 8). Die entsprechenden Zitate finden sich in Abschnitt A6.1. Alle diese Studien konnten aus öffentlichen Quellen identifiziert werden.

Tabelle 8: Studienpool

Studie	Verfügbare Dokumente		
	Vollpublikation (in öffentlich zugänglichen Fachzeitschriften)	Ergebnisbericht aus Studienregistern	Studienbericht aus Herstellerunterlagen (nicht öffentlich zugänglich)
Vergleichende Interventionsstudien der Screeningkette			
Kubiak 2014	[22]	nein	entfällt
Vergleichende Interventionsstudien zum Therapiebeginn			
Brown 2011	[24]	nein	entfällt
Myers 2002	[25]	nein	entfällt
Cuvelier 2016	[26]	nein	entfällt
Vergleichende Studien zur diagnostischen Güte			
Kwan 2015	[27]	nein ^a	entfällt
Vogel 2014	[28]	nein	entfällt
Kwan 2013	[23]	nein	entfällt
Verbsky 2012	[29]	nein	entfällt
Routes 2009	[17]	nein	entfällt
Chien 2015	[30]	nein	entfällt
Kanegae 2016	[31]	nein	entfällt
de Felipe 2016	[32]	nein	entfällt
a: bezieht sich nur auf einen Teil der publizierten Daten, der vorliegend nicht berücksichtigt wird (Daten aus Pilotstudie)			

A3.1.4 Studien unklarer Relevanz

Von den in Studienregistern identifizierten Studien unklarer Relevanz lässt 1 laufende Studie aus Frankreich Ergebnisse als vergleichende Interventionsstudie der Screeningkette erwarten [34]. Als Studie zur Kosteneffizienz eines Screenings auf SCID sollen in die Interventionsgruppe 200 000 Neugeborene an 50 Geburtskliniken eingeschlossen werden, die innerhalb der ersten 72 Lebensstunden mittels Filterkartenblut anhand des Markers TREC auf SCID getestet werden sollen. Als Kontrollgruppe dienen diejenigen mit SCID diagnostizierten aus 700 000 Neugeborenen in Frankreich, die kein Screening erhalten haben. Als patientenrelevante Endpunkte sind unter anderem Ergebnisse zur Mortalität sowie Morbidität (Infektionen) angekündigt. Für positiv getestete Kinder der Interventionsgruppe ist als Referenztest eine Lymphozytensubgruppenanalyse vorgesehen. Kinder, die nicht mittels SCID-Screening identifiziert wurden, sollen mittels Registeranalyse nachbeobachtet werden. Damit sind auch Daten zur diagnostischen Güte (Sensitivität und Spezifität) zu erwarten. Bis Mitte 2018 sollen Studiendaten vorliegen.

A3.2 Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudien der Screeningkette

A3.2.1 Studiendesign und Studienpopulationen

In der Studie **Kubiak 2014** [22] wurden Daten zu Neugeborenen mit SCID in 3 US-amerikanischen Krankenhäusern in unterschiedlichen Staaten präsentiert. Die Daten umfassen neben Angaben zum Zeitpunkt des Therapiebeginns die Art der Diagnose (Neugeborenen-screening auf SCID oder kein Screening) sowie Angaben zu den patientenrelevanten Endpunkten (Mortalität, Morbidität, Hospitalisierung vor Diagnose). Im Studienzeitraum November 2005 bis November 2011 umfasste das Neugeborenen-screening in Kalifornien seit August 2010 ein SCID-Screening [23]. Somit konnte die Studie auch als vergleichende Studie der Gruppen Screening versus kein Screening interpretiert werden.

In den folgenden Tabellen sind die Studien- und Interventionscharakteristika, die wesentlichen Ein- und Ausschlusskriterien sowie die Studienpopulation beschrieben.

Tabelle 9: Charakterisierung der eingeschlossenen Studie – Studie zur Screeningkette

Studie	Studien- design	Personen- zahl N	Inter- vention Zahl der Personen	Vergleich Zahl der Personen	Ort und Zeitraum der Rekrutierung	Studien- dauer	Relevante Endpunkte
Kubiak 2014 ^a	retro- spektive Kohorten- analyse (ökono- mische Analyse)	20 ^b	2	18	USA: Kalifornien ^c , Florida ^d und Pennsylvanien ^d	November 2005 bis November 2011	Mortalität vor Transplantation, Morbidität (Kompli- kationen), Hospitalisierung vor Diagnose
<p>a: Ergebnisse werden nicht dargestellt. b: Von 25 eingeschlossenen Patienten haben 5 anstelle der Stammzelltransplantation eine Gentherapie erhalten. c: Einführung des SCID-Screenings in 2010 [23]. d: Ohne SCID-Screening zur Zeit der Studienlaufzeit. k. A.: keine Angabe; N: Anzahl eingeschlossener Patienten; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)</p>							

Tabelle 10: Ein- / Ausschlusskriterien für Personen in der Studie – Studie zur Screeningkette

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Kubiak 2014	Neugeborene mit SCID in 3 teilnehmenden Zentren in 3 Staaten der USA (Kalifornien, Florida und Pennsylvanien)	k. A.
k. A.: keine Angabe; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)		

Tabelle 11: Charakterisierung der Studienpopulationen – Studie zur Screeningkette

Studie	N	Geschlecht w / m	Alter	Prävalenz
Kubiak 2014				
Interventionsgruppe	2	k. A.	Neugeborene	- ^a
Vergleichsgruppe	13	k. A.	Neugeborene	- ^a
a: In der Studie sind nur die SCID-Fälle berichtet, nicht die Gesamtpopulation. k. A.: keine Angabe; m: männlich; N: Anzahl eingeschlossener Patienten; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); w: weiblich				

Tabelle 12: Charakterisierung der Intervention

Studie	Berichtete Screeningstrategie	Nachbeobachtung
Kubiak 2014	k. A.	k. A.
k. A.: keine Angabe		

A3.2.2 Einschätzung des Verzerrungspotenzials auf Studienebene

Die Einschätzung des Verzerrungspotenzials auf Studienebene entfällt, da die Ergebnisse der Studie nicht dargestellt werden (siehe Abschnitt A3.3).

A3.3 Patientenrelevante Endpunkte

Da die Ergebnisse der Studie nicht interpretierbar sind, entfällt ihre Darstellung ebenso wie die Darstellung ihres Verzerrungspotenzials.

Der maßgebliche Grund dafür, dass die Studie nicht als vergleichende Interventionsstudie der Screeningkette ausgewertet wurde, liegt darin, dass mangels Angabe der Nachbeobachtungsdauer der Gruppenunterschied in den berichteten Endpunkten nicht auf einen Zeitraum bezogen werden konnte. Damit ist die Vergleichbarkeit der Nachbeobachtung in den Gruppen unklar. Der angegebene Studienzeitraum ermöglicht keine Vergleichbarkeit und spricht im Gegenteil dafür, dass die Nachbeobachtungszeit der im Screening gefundenen Neugeborenen deutlich kürzer ist als in der Vergleichsgruppe, da das Screening erst 2010 und damit gegen Ende der Studienlaufzeit (2011) eingeführt wurde. Ein weiterer Grund dafür, die Ergebnisse dieser Studie nicht darzustellen, liegt darin, dass die Ergebnisse für die berichteten

Endpunkte (abgesehen von Mortalität) nicht eindeutig den Gruppen Screening / kein Screening zugeordnet werden können.

A3.4 Charakteristika der in die Bewertung eingeschlossenen vergleichenden Interventionsstudien zum Therapiebeginn

A3.4.1 Studiendesign und Studienpopulationen – Studien zum Therapiebeginn

In der retrospektiven Studie **Brown 2011** aus dem Vereinigten Königreich [24] – bestand die Interventionsgruppe aus 60 Kindern, die aufgrund einer positiven Familienhistorie, meistens aufgrund eines betroffenen Geschwisterkindes, pränatal oder bei der Geburt getestet wurden und bei denen SCID diagnostiziert wurde. Die 48 betroffenen Geschwisterkinder respektive Familienmitglieder, die erst nach Präsentation mit auffälliger Symptomatik diagnostiziert worden waren, bildeten die Vergleichsgruppe. Angaben zur Nachbeobachtungsdauer wie auch zum Transplantationszeitpunkt fehlen. Beschrieben ist aber, dass die Kinder der Interventionsgruppe eine infektionsprophylaktische Therapie erhielten, die Kinder der Vergleichsgruppe dagegen nicht. Es werden die Ergebnisse zu den Endpunkten Mortalität (vor und nach Transplantation) und Morbidität berichtet, wobei diese nicht auf eine frühere kurative Therapie, sondern auf die infektionsprophylaktische Therapie zu beziehen sind.

Im Unterschied zur Studie von Brown diente in der retrospektiven / prospektiven Studie **Myers 2002** [25] aus den USA der Zeitpunkt der Transplantation als Merkmal der Gruppenzuteilung. Die Interventionsgruppe bestand aus Kindern, die aufgrund einer positiven Familienhistorie pränatal oder bei der Geburt diagnostiziert und vor dem 28. Lebensjahr transplantiert wurden (n = 21). Für die 96 nach dem 28. Lebensjahr transplantierten Kinder der Vergleichsgruppe fehlt die explizite Angabe des Diagnosezeitpunktes. Die Autoren bezogen in die Ergebnisanalysen aus dieser Gruppe die 71 überlebenden Kinder ein.

In der vergleichenden retrospektiven Kohortenstudie **Cuvelier 2016** [26] bestand die Interventionsgruppe aus 2 Kindern mit ZAP70-Defizienz, einer seltenen autosomal-rezessiven SCID-Form, die aufgrund positiver Familienhistorie asymptomatisch identifiziert und innerhalb der ersten beiden Lebensmonate transplantiert wurden. Die Vergleichsgruppe bildeten 6 Kinder, ebenfalls mit ZAP70-Defizienz. In keiner der beiden Gruppen gab es einen Todesfall. Der einzig berichtete Todesfall bezieht sich auf ein Kind, das wegen fehlender abschließender Diagnose einer ZAP70-Defizienz aus der Analyse ausgeschlossen wurde. Auf eine Darstellung der Ergebnisse zur Mortalität wird daher verzichtet. Aus dieser Studie werden die Daten zur Morbidität (Infektionen zum Zeitpunkt der Transplantation) berichtet. Da nur insgesamt 8 Kinder in dieser Studie untersucht wurden, kann sie in Bezug auf Endpunkte, zu denen keine weiteren Studien Daten berichtet haben, keine belastbare Datengrundlage für die Nutzenbewertung beitragen. Auf die Darstellung weiterer in der Studie berichteter Endpunkte zur Morbidität (zum Beispiel GVHD, Anzahl Transplantationen) wird daher verzichtet.

Tabelle 13: Charakterisierung der eingeschlossenen Studien – Studien zum Therapiebeginn

Studie	Studien- design	Patienten- zahl	Intervention Zahl der Patienten	Vergleich Zahl der Patienten	Ort und Zeitraum der Durchführung	Studien- dauer	Relevante Endpunkte ^a
Brown 2011	retrospektive vergleichende Kohorten- studie	108	60 ^{b,c}	48 ^d	Vereinigtes Königreich, 1979–2010 ^e	k. A.	Mortalität Morbidität (Infektionen) ^f
Myers 2002	retrospektive / prospektive vergleichende Kohorten- studie ^g	117	21	96	USA, North Carolina, Durham k. A. ^h	k. A.	Mortalität ⁱ
Cuvelier 2016	retrospektive vergleichende Kohorten- studie	9 ^j	2 ^k	6	Canada, Winnipeg, 1992–2014 ^l	k. A.	Mortalität ^m Morbidität (Infektionen)

a: In den Publikationen wurde nicht zwischen primären und sekundären Endpunkten unterschieden. Hier wurden ausschließlich Endpunkte angegeben, die für diese Nutzenbewertung relevant waren.

b: Der Zeitpunkt der Intervention (Stammzelltherapie) ist nicht angegeben. Der Zeitpunkt der Diagnose ist mit 0 im Median angegeben; die Spanne umfasst 0–29 Lebensstage. Die Gruppeneinteilung erfolgte aufgrund der Kennzeichnung des Diagnosezeitpunktes als pränatal oder bei der Geburt und der Angabe der Durchführung prophylaktischer Maßnahmen vor Transplantation bei allen Neugeborenen dieser Gruppe.

c: Die Studienpopulation von 60 Patienten umfasste auch 1 Patienten ohne Stammzelltransplantation (stattdessen Gentherapie).

d: Der Zeitpunkt der Intervention (Stammzelltherapie) ist nicht angegeben. Der Zeitpunkt der Diagnose ist mit 143,5 Tagen im Median angegeben; die Spanne umfasst 1–455 Tage. Die Gruppeneinteilung erfolgte aufgrund der separierenden Beschreibung in der Studie sowie der Angabe, dass prophylaktische Maßnahmen vor Transplantation nicht durchgeführt wurden.

e: Interventionsgruppe: 1982–2010; Vergleichsgruppe: 1979–2009.

f: Für den Endpunkt Morbidität sind Daten für die Vergleichsgruppe nur von 29 Patienten berichtet (60 %). Für die Interventionsgruppe liegen Daten von 57 von 60 Patienten (95 %) vor.

g: Aus der Publikation geht nicht hervor, welcher Teil prospektiv durchgeführt wurde.

h: Die Patienten wurden „während der letzten 19,2 Jahre“ im betreffenden Krankenhaus behandelt.

i: In der Auswertung wurden nur Endpunkte für die Überlebenden nach der Transplantation in der Vergleichsgruppe berichtet, daher sind ausschließlich Daten zur Mortalität für die Nutzenbewertung verwendbar.

j: Patienten von 6 mennonitischen Familien mit Zeta-Chain-associated Protein 70 kDa Deficiency (ZAP70-Defizienz), 1 Kind davon wurde aus der Analyse ausgeschlossen.

k: Patienten, die asymptomatisch aufgrund positiver Familienhistorie identifiziert wurden; Zeitpunkt der Transplantation: 1 und 2 Monate

l: Zeitraum, in dem die Transplantationen durchgeführt wurden

m: Die Autoren haben das einzige verstorbene Kind aus der Analyse ausgeschlossen, Bei diesem Kind wurde das Vorliegen eines SCID nicht abschließend diagnostiziert und eine Zuordnung zu einer der Gruppen ist nicht möglich. Daher waren die Daten zur Mortalität nicht interpretierbar und es sind ausschließlich die Daten zur Morbidität für die Nutzenbewertung verwendbar.

k. A.: keine Angabe

Tabelle 14: Ein- / Ausschlusskriterien für Patienten – Studien zum Therapiebeginn

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Brown 2011	<ul style="list-style-type: none"> ▪ Kinder mit SCID ▪ Stammzelltransplantation im Great Ormond Street Hospital National Health Service Trust und Newcastle General Hospital zwischen 1979 und 2010 <ul style="list-style-type: none"> ▫ als SCID-Patient mit positiver Familienhistorie (Interventionsgruppe) ▫ als vorangegangener erstpräsentierender Familienangehöriger mit SCID (Vergleichsgruppe) 	k. A.
Myers 2002	<ul style="list-style-type: none"> ▪ Kinder mit SCID ▪ Stammzelltransplantation im Duke University Medical Center in einem Zeitraum von 19,2 Jahren 	k. A.
Cuvelier 2016	<ul style="list-style-type: none"> ▪ Kinder mit ZAP70-Defizienz Stammzelltransplantation im Rahmen des Manitoba Blood and Marrow Transplant Program seit 1992 ▪ als SCID-Patient mit positiver Familienhistorie (Interventionsgruppe) ▪ als erstpräsentierender SCID-Patient ohne positive Familienhistorie (Vergleichsgruppe) 	k. A.
k. A.: keine Angabe; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); ZAP70: Zeta-Chain-associated Protein 70 kDa		

Tabelle 15: Charakterisierung der Studienpopulationen – Studien zum Therapiebeginn

Studie	N	Geschlecht w / m	Alter bei Diagnose [Tage], Median [Min; Max]	Alter bei Transplantation [Tage], Median [Min; Max]
Brown 2011	108			
Interventionsgruppe	60 ^a	k. A.	0 [0; 29]	k. A.
Vergleichsgruppe	48	k. A.	143,5 [1; 455]	k. A.
Myers 2002	117 ^b			
Interventionsgruppe	21	2 / 19	k. A. ^c	10 [7; 24]
Vergleichsgruppe	71 ^d	16 / 55 ^d	k. A.	190 [45; 516] ^d
Cuvelier 2016	8			
Interventionsgruppe	2	1/1	3 Wochen bzw. 12 Tage ^e	1 bzw. 2 Monate ^f
Vergleichsgruppe	6	5/1	10 [4; 27] Monate	11,5 [5,5; 43] Monate
a: Die Studienpopulation von 60 Patienten umfasste auch 1 Patienten ohne Stammzelltransplantation (stattdessen Genterapie). b: Demografische Angaben zu den vor der Stammzelltransplantation Verstorbenen (n = 25) fehlen. c: Diagnose in utero (n = 9) oder bei Geburt (n = 12); die Angabe der Gestationswoche / des Lebensstages fehlt. d: Diese Angabe bezieht sich auf die überlebenden Patienten. e: Diagnosezeitpunkt der beiden Patienten f: Transplantationszeitpunkt der beiden Patienten k. A.: keine Angabe; m: männlich; N: Anzahl eingeschlossener Patienten; w: weiblich				

Tabelle 16: Charakterisierung der Interventionen – Studien zum Therapiebeginn

Studie	Intervention	Vergleich
Brown 2011	Stammzelltransplantation ^a	Stammzelltransplantation
	<ul style="list-style-type: none"> ▪ Diagnose: pränatal oder bei Geburt ▪ prophylaktische Maßnahmen^b: Antibiotika; Immunglobulinersatztherapie 	<ul style="list-style-type: none"> ▪ Diagnose: postnatal^c ▪ prophylaktische Maßnahmen: keine
Myers 2002	Stammzelltransplantation	Stammzelltransplantation
	<ul style="list-style-type: none"> ▪ Alter bei Transplantation: ≤ 28 Lebensstage ▪ ohne Konditionierung ▪ ohne Graft-versus-Host Reaktion-Prophylaxe ▪ Spendereigenschaften: HLA-identische Spender; HLA-haploidentische Spender 	<ul style="list-style-type: none"> ▪ Alter bei Transplantation: > 28 Lebensstage ▪ allogene
Cuvelier 2016	Stammzelltransplantation	Stammzelltransplantation
	<ul style="list-style-type: none"> ▪ Diagnose: postnatal in asymptomatischem Zustand ▪ prophylaktische Maßnahmen: k. A. 	<ul style="list-style-type: none"> ▪ Diagnose: postnatal ▪ prophylaktische Maßnahmen: k. A.
<ul style="list-style-type: none"> ▪ ohne Konditionierung, reduzierte Konditionierung oder myeloablative Konditionierung ▪ Spendereigenschaften: Geschwisteranteil („matched sibling donor“); Familienmitglied („matched family donor“); nicht verwandter Spender („matched / mismatched unrelated donor“) 		
<p>a: Die Studienpopulation von 60 Patienten umfasste auch 1 Patienten ohne Stammzelltransplantation (stattdessen Genterapie).</p> <p>b: Der Zeitpunkt des Behandlungsbeginns (Lebensstag) ist nicht angegeben.</p> <p>c: Es ist nicht angegeben, ob die Diagnose aufgrund klinischer Präsentation oder bei asymptomatischen Kindern aus anderen Gründen wie zum Beispiel aufgrund einer auffälligen Familienhistorie erfolgte.</p> <p>HLA: Human Leukocyte Antigen (menschliches Leukozytenantigen); k. A.: keine Angabe</p>		

A3.4.2 Einschätzung des Verzerrungspotenzials auf Studienebene – Studien zum Therapiebeginn

Die Einschätzung des Verzerrungspotenzials auf Studienebene ist in der folgenden Tabelle 17 dargestellt. Die 3 Studien wurden als potenziell hochverzerrt beurteilt. Für die Studie **Brown 2011** wurde eine Autorenanfrage gestellt, um den zeitlichen Versatz der verglichenen Gruppen sowie die Dauer der Nachbeobachtung zu klären. Die Anfrage wurde nicht inhaltlich

beantwortet. Die Studie **Myers 2002** weist dieselben Unklarheiten auf. Die Autoren wurden jedoch nicht zur Klärung angeschrieben, da die Ergebnisse dieser Studie keinen dramatischen Effekt zeigen (Abschnitt A3.5.2).

Tabelle 17: Verzerrungspotenzial auf Studienebene – Studien zum Therapiebeginn

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung prognostisch relevanter Faktoren	Verblindung					Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte		
Brown 2011	unklar ^a	unklar ^b	nein ^c	nein ^c	unklar ^d	ja	hoch ^e	
Myers 2002	unklar ^f	unklar ^b	nein ^g	nein ^g	unklar ^{d,h}	ja	hoch ^e	
Cuvelier 2016	unklar ⁱ	unklar ^b	nein ^c	nein ^c	unklar ^j	ja	hoch ^e	

a: Die Patienten der Interventions- und Kontrollgruppe wurden zeitlich leicht versetzt eingeschlossen.
 b: In der Publikation finden sich keine ausreichenden Angaben zum zeitlichen Versatz.
 c: In der Publikation finden sich keine Angaben. Die Daten wurden retrospektiv aus der Praxis erhoben, daher ist von fehlender Verblindung auszugehen.
 d: Es fehlen Angaben zur Fallzahlplanung, zur Nachbeobachtungsdauer und zur Vollständigkeit der Daten.
 e: Es handelt sich um eine nicht randomisierte Studie.
 f: Es fehlt die Angabe der Geburtsjahre für die Kontrollgruppe.
 g: In der Publikation finden sich keine Angaben. Die Studie wird als retrospektiv / prospektiv angegeben. Unklar ist, auf welchen Teil der Studie sich die prospektive Vorgehensweise bezieht. Jedenfalls für den retrospektiven Teil ist von fehlender Verblindung auszugehen.
 h: Die Ergebnisberichterstattung erfolgt nur für Interventionsgruppe bzgl. Überlebensdauer seit Transplantation und Morbidität und damit einseitig.
 i: Die Patienten der Interventions- und Kontrollgruppe wurden ab 1992 eingeschlossen; die Geburtsjahre sind nicht angegeben. Somit ist eine unterschiedliche Dauer der Follow-ups möglich.
 j: Keine Angabe der geplanten Auswertungen. Ausschluss eines vor Stammzelltransplantation verstorbenen Kindes mit fehlender Diagnose, aber starken Hinweisen darauf, dass es ZAP70 hatte, ohne Zuordnung zu Interventions- oder Kontrollgruppe (einziger Todesfall).
 ZAP70: Zeta-Chain-associated Protein 70 kDa

A3.5 Patientenrelevante Endpunkte aus vergleichenden Interventionsstudien zum Therapiebeginn

Für die Nutzenbewertung wurden Daten aus 2 Studien zu dem Endpunkt Mortalität sowie Daten aus 2 Studien zum Endpunkt Morbidität (Infektionen) herangezogen.

Tabelle 18: Übersicht zur Extraktion von patientenrelevanten Endpunkten, Datenverfügbarkeit

Endpunkt	Morbidity							Gesundheitsbezogene Lebensqualität sowie psychosoziale Aspekte
	Mortalität	Infektionen	Graft-versus-Host-Disease	Erneute Transplantation	Krankenhausaufenthalte	Entwicklungsstörungen	Unerwünschte Ereignisse	
Studie								
Brown 2011	●	○	-	-	-	-	-	-
Myers 2002	●	-	-	-	-	-	-	-
Cuvelier 2016	-	○	(○)	(○)	-	-	-	-

-: keine Daten vorhanden; (-): Daten nicht anwendbar; ○: Daten vorhanden und anwendbar; ●: Daten in einer Meta-Analyse berücksichtigt; (○): Da nur insgesamt 8 Kinder in dieser Studie untersucht wurden, kann sie in Bezug auf Endpunkte, zu denen keine weiteren Studien Daten berichtet haben, keine belastbare Datengrundlage für die Nutzenbewertung beitragen. Auf die Darstellung wird daher verzichtet.

A3.5.1 Verzerrungspotenzial auf Endpunktebene – Mortalität und Morbidity

Das Verzerrungspotenzial aller Endpunkte ist in Tabelle 19 und Tabelle 20 dargestellt. Das hohe Verzerrungspotenzial auf Studienebene schlägt sich direkt nieder auf das endpunktspezifische Verzerrungspotenzial. Die Ergebnisse zur Mortalität und Morbidity (Infektionen) wurden als hochverzerrt beurteilt.

Tabelle 19: Verzerrungspotenzial auf Endpunktebene – Mortalität

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Endpunktebene
Brown 2011	hoch ^a	nein ^b	unklar ^c	unklar ^d	ja	hoch ^e
Myers 2002	hoch ^a	nein ^f	unklar ^c	unklar ^{d,g}	ja	hoch ^e

a: Es handelt sich um eine nicht randomisierte Studie.
b: In der Publikation finden sich keine Angaben. Die Daten wurden retrospektiv aus der Praxis erhoben, daher ist von fehlender Verblindung auszugehen.
c: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-Follow-up-Patienten nicht einschätzbar.
d: Es fehlen Angaben zur Fallzahlplanung und zur Nachbeobachtungsdauer.
e: Das hohe Verzerrungspotenzial auf Studienebene schlägt sich direkt auf das Verzerrungspotenzial auf Endpunktebene nieder.
f: In der Publikation finden sich keine Angaben. Die Studie wird als retrospektiv / prospektiv angegeben. Unklar ist, auf welchen Teil der Studie sich die prospektive Vorgehensweise bezieht. Jedenfalls für den retrospektiven Teil ist von fehlender Verblindung auszugehen.
g: Angaben zur Überlebensdauer seit Transplantation nur für Interventionsgruppe.
ITT: Intention to treat

Tabelle 20: Verzerrungspotenzial auf Endpunktebene – Morbidität

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Endpunktebene
Brown 2011	hoch ^a	nein ^b	nein ^c	unklar ^d	ja	hoch ^e
Cuvelier	hoch ^a	nein ^b	nein ^f	ja	ja	hoch ^e

a: Es handelt sich um eine nicht randomisierte Studie.
b: In der Publikation finden sich keine Angaben. Die Daten wurden retrospektiv aus der Praxis erhoben, daher ist von fehlender Verblindung auszugehen.
c: Für 3 von 60 Patienten der Interventionsgruppe und für 19 von 48 Patienten der Kontrollgruppe wurden die Daten als „nicht verfügbar“ angegeben.
d: Es fehlen Angaben zur Fallzahlplanung und zur Nachbeobachtungsdauer.
e: Das hohe Verzerrungspotenzial auf Studienebene schlägt sich direkt nieder auf das Verzerrungspotenzial auf Endpunktebene.
f: Ausschluss eines vor Stammzelltransplantation verstorbenen Kindes.
ITT: Intention to treat; ZAP70: Zeta-Chain-associated Protein 70 kDa

A3.5.2 Ergebnisse zum Endpunkt Mortalität

Die Ergebnisse der relevanten Studien zum Endpunkt Mortalität sind in Tabelle 21 abgebildet.

Die Ergebnisse zur Gesamtmortalität aus der Studie **Brown 2011** [24] zeigen mit 10 % zu 60 % einen dramatischen Effekt zugunsten des früheren Therapiebeginns im Sinne eines früheren Behandlungsbeginns mit Infektionsprophylaxe (nicht: Transplantationszeitpunkt; der ist in dieser Studie nicht berichtet) im Vergleich zu einem späteren Therapiebeginn. Noch deutlicher ist der Effekt beim Vergleich der Mortalität zum Zeitpunkt vor der Transplantation mit 1,8 % in der Interventions- im Vergleich zu 35 % in der Vergleichsgruppe.

Für den Zeitpunkt nach der Transplantation zeigen die Ergebnisse mit 8,5 % in der Interventions- und 39 % in der Vergleichsgruppe auch einen deutlichen Vorteil.

Für den Endpunkt Mortalität liegt ein Anhaltspunkt für einen Nutzen eines früheren Beginns mit einer infektionsprophylaktischen Therapie, die in eine kurative Therapie mündet, im Vergleich zu einem späteren Beginn vor.

Die Studie **Myers 2002** [25] zeigt einen statistisch nicht signifikanten Vorteil eines früheren versus späteren Behandlungsbeginns, bezogen auf die kurative Therapie.

Der gepoolte Gesamteffekt beider Studien zeigt einen statistisch signifikanten Vorteil eines früheren Therapiebeginns, wobei diese Gesamtbetrachtung außer Acht lässt, dass in beiden Studien der Beginn der unterschiedlichen Therapien (Zeitpunkt der Infektionsprophylaxe einerseits und Zeitpunkt der Stammzelltransplantation andererseits) untersucht wurde.

Tabelle 21: Ergebnisse – Mortalität

Studie	Endpunkt Definition	Intervention		Vergleich		Früherer versus späterer Therapiebeginn OR [95 %-KI] ^b ; p-Wert ^c
		N	Patienten mit Ereignissen n (%)	N	Patienten mit Ereignissen n (%)	
Brown 2011	Gesamtmortalität ^a	60	6 (10)	48	29 (60)	0,07 [0,03; 0,20]; < 0,001
	davon Mortalität vor Transplantation	60	1 (1,8)	48	17 (35)	0,03 [0,00; 0,24]; < 0,001
	davon Mortalität nach Transplantation	60	5 (8,5)	48	12 (39)	0,27 [0,09; 0,84]; 0,019
Myers 2002	Gesamtmortalität ^a	21	1 (5)	96	25 (26)	0,14 [0,02; 1,11]; 0,080

a: keine Angabe des Zeitraums, auf den sich die Angaben beziehen
b: eigene Berechnung
c: eigene Berechnung, unbedingter exakter Test (CSZ-Methode nach [63])
k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl ausgewerteter Patienten; n: Anzahl Patienten in Gruppe; OR: Odds Ratio

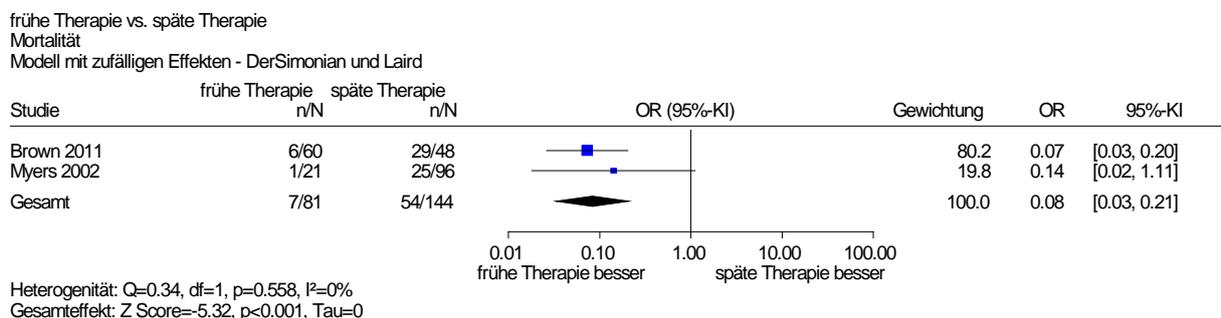


Abbildung 2: Gepoolter Gesamteffekt Therapiebeginn

A3.5.3 Ergebnisse zum Endpunkt Morbidität (Infektionen)

Zum Endpunkt Morbidität (Infektionen) liegen Daten aus 2 Studien vor. Die Daten aus der Studie **Brown 2011** [24] (der konkrete Zeitpunkt der Infektionen ist nicht angegeben) sind unvollständig. Eine Autorenanfrage mit der Bitte um eine Begründung dieser Diskrepanz wurde nicht inhaltlich beantwortet (Abschnitt A3.1.2.4).

Um die Daten dennoch zu verwerten, wurde eine eigene Sensitivitätsanalyse gerechnet, in der die Ereignisse auf die vollständigen Gruppen und nicht nur die tatsächlich nachbeobachteten Patienten bezogen wurden. Auch diese (konservative) Schätzung zeigt einen statistisch

signifikanten Gruppenunterschied. Dies stützt die Ergebnisse zur Mortalität. Der Effekt ist zwar groß, er könnte aber eventuell allein auf Verzerrung zurückzuführen sein.

Aus der Studie **Cuvelier 2016** [26] wurden die Daten zu Infektionen zum Zeitpunkt der Stammzelltransplantation herangezogen. Sie zeigen einen statistisch nicht signifikanten Gruppenunterschied. Da in den beiden Studien der Zeitpunkt des Vorliegens der Infektionen nicht übereinstimmt, wurde keine Meta-Analyse durchgeführt.

Tabelle 22: Ergebnisse – Morbidität

Studie	Endpunkt Definition	Intervention		Vergleich		Früherer vs. späterer Therapiebeginn OR [95 %-KI] ^d ; p-Wert ^e
		N	Patienten mit Ereignissen n	N	Patienten mit Ereignissen n	
Brown 2011	Infektionen ^a	57 ^b	10	29 ^c	26	
Sensitivitäts- analyse		60	10	48	26	0,17 [0,07; 0,41]; < 0,001
Cuvelier 2016	Infektionen zum Zeitpunkt der Stammzell- transplantation	2	0	6	5	0,05 [0; 1,87]; 0,055
<p>a: Der konkrete Zeitpunkt der Infektionen ist nicht angegeben. b: Für 3 von 60 Patienten wurden die Daten als „nicht verfügbar“ angegeben. c: Für 19 von 48 Patienten wurden die Daten als „nicht verfügbar“ angegeben. d: Eigene Berechnung. e: Eigene Berechnung, unbedingter exakter Test (CSZ-Methode nach [63]). k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl ausgewerteter Patienten; n: Anzahl Patienten in Gruppe</p>						

Für den Endpunkt Morbidität (Infektionen) liegt kein Anhaltspunkt für einen (Zusatz-)Nutzen oder (höheren) Schaden vor.

A3.5.4 Subgruppenmerkmale und andere Effektmodifikatoren

Subgruppenmerkmale und andere Effektmodifikatoren wurden nicht untersucht, da keine ausreichenden Daten vorlagen.

A3.6 Charakteristika der in die Bewertung eingeschlossenen Studien zur diagnostischen Güte

A3.6.1 Studiendesign und Studienpopulationen

In den folgenden Tabellen sind die eingeschlossenen Studien charakterisiert.

Von den 8 Studien berichten 4 die Ergebnisse des um ein SCID-Screening erweiterten Neugeborenen Screenings in Staaten der USA. In 1 weiteren Studie stammen die Daten dagegen aus 2 Krankenhäusern in Navajo Nation, USA; die Daten aus Taiwan wurden in 1 Krankenhaus erhoben (Tabelle 23 und Tabelle 24). Hinzu kommen 1 Studie aus Brasilien, die Daten von Neugeborenen aus 2 Krankenhäusern berichtet [31], sowie 1 Studie aus Spanien mit Daten aus 3 Krankenhäusern [32].

Die Studienpopulation besteht jeweils aus Neugeborenen. Auch wenn Angaben zu demografischen Daten im Wesentlichen fehlen, kann davon ausgegangen werden, dass sie wegen der Testung im Rahmen des etablierten Neugeborenen Screenings, abgesehen von der Prävalenz (dazu siehe Abschnitt A3.6.2.2), keine maßgeblichen Unterschiede aufweisen.

Die Screeningalgorithmen unterscheiden sich im Detail in Hinblick auf die Trennwerte sowie auf den Einsatz von Kontrollmarkern zur Prüfung der Probenqualität. Übereinstimmend wurden aber jeweils Proben aus beim Neugeborenen Screening gewonnenem Filterkartenblut genommen, um diese auf die Anzahl von TRECs als Marker für SCID zu testen. Bei den Referenztests bestehen wiederum Unterschiede (Tabelle 26).

Tabelle 23: Charakterisierung der eingeschlossenen Studien – Studien zur Testgüte

Studie	Studiendesign	Patientenanzahl	Studienziel	Evidenzstufe ^a	Ort und Zeitraum der Durchführung	Zielgrößen
Kwan 2015	prospektive Kohortenstudie, VOPT	6100	Beschreibung der Ergebnisse eines SCID-Screenings in Navajo Nation ^b	I Ib	USA: Navajo Nation, Februar 2012 bis Juli 2014	falsch-positive, richtig-positive Testergebnisse ^c
Vogel 2014 ^d	prospektive Kohortenstudie, VOPT	485 912	Beschreibung der Ergebnisse eines SCID-Screenings in New York State ^e	I Ib	USA: New York State, 29. September 2010 bis 28. September 2012	falsch-positive, richtig-positive Testergebnisse ^c
Kwan 2013	prospektive Kohortenstudie, VOPT	993 724	Beschreibung der Ergebnisse eines SCID-Screenings in Kalifornien	I Ib	USA: Kalifornien, August 2010 bis 2012 ^f	falsch-positive, richtig-positive Testergebnisse ^c
Verbsky 2012 ^d	prospektive Kohortenstudie, VOPT ^h	207 696	Beschreibung der Ergebnisse eines SCID-Screenings in Wisconsin ^g	I Ib	USA: Wisconsin, 1. Januar 2008 bis 31. Dezember 2010	falsch-positive, richtig-positive Testergebnisse, PPV ^c
Routes 2009 ^d	prospektive Kohortenstudie, VOPT	71 000	Beschreibung der Ergebnisse eines SCID-Screenings in Wisconsin	I Ib	USA: Wisconsin, 1. Januar bis 31. Dezember 2008	falsch-positive, richtig-positive Testergebnisse ^{c,h}
Chien 2015	prospektive Kohortenstudie, VOPT	106 391	Inzidenz von SCID	I Ib	Taiwan, 1. Mai 2010 bis 31. Dezember 2011	falsch-positive, richtig-positive Testergebnisse ^c
Kanegae 2016	prospektive Kohortenstudie, VOPT	8682	Beschreibung der Ergebnisse eines SCID-Screenings in Brasilien	I Ib	Brasilien, São Paulo und Pará; k. A. zum Zeitraum	falsch-positive, richtig-positive Testergebnisse ^c
de Felipe 2016	prospektive Kohortenstudie, VOPT	3810 ⁱ	Beschreibung der Ergebnisse eines SCID-Screenings in Spanien, Sevilla	I Ib	Spanien: Sevilla, Februar 2014 bis März 2015	falsch-positive, richtig-positive Testergebnisse ^c

(Fortsetzung)

Tabelle 23: Charakterisierung der eingeschlossenen Studien – Studien zur Testgüte
(Fortsetzung)

<p>a: Entspricht der Evidenzstufe des G-BA.</p> <p>b: Es werden Ergebnisse einer Pilotstudie sowie Ergebnisse seit der Einführung des SCID-Screenings beschrieben. Die Ergebnisse der Pilotstudie wurden für die Bewertung der diagnostischen Güte nicht berücksichtigt.</p> <p>c: Es werden keine Zielgrößen definiert, aber falsch-positive und richtig-positive Testergebnisse angegeben oder aber es liegen Angaben vor, die eine entsprechende Einordnung erlauben.</p> <p>d: Ergebnisse werden nicht dargestellt</p> <p>e: Es werden Ergebnisse seit der Einführung des SCID-Screenings beschrieben, wobei der Trennwert für den SCID-Test im Studienverlauf verändert wurde.</p> <p>f: Es werden die Ergebnisse der „ersten 2 Jahre“ berichtet, der Zeitraum ist nicht konkret angegeben.</p> <p>g: Die Daten von Routes 2009 sind in den von Verbsky 2012 berichteten Ergebnissen enthalten. Die angewandten Screeningalgorithmen weichen leicht voneinander ab; ein Hinweis hierauf fehlt bei Verbsky 2012.</p> <p>h: K. A. zur Nachbeobachtung; es findet sich allein der Hinweis, dass es im Studienzeitraum am Studienort keine SCID-Fälle gab.</p> <p>i: Es werden nur die um die Daten der Pilotstudie [33] bereinigten Daten berichtet (3810 von 5160); aus der Autorenantwort geht eine Anzahl von 5163 eingeschlossenen Neugeborenen hervor.</p> <p>G-BA: Gemeinsamer Bundesausschuss; k. A.: keine Angabe; PPV: Positive predictive Value (positiver prädiktiver Wert); SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); VOPT: Verification of only positive Testers</p>

Tabelle 24: Ein- / Ausschlusskriterien für Personen in den Studien – Studien zur Testgüte

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Kwan 2015	Neugeborene in 2 Krankenhäusern in Chinle und Tuba City, Arizona, USA	k. A.
Vogel 2014 ^a	Neugeborene in New York State, USA	k. A.
Kwan 2013	Neugeborene in Kalifornien, USA	k. A.
Verbsky 2012 ^a	Neugeborene in Wisconsin, USA	k. A.
Routes 2009 ^a	Neugeborene in Wisconsin, USA	k. A.
Chien 2015	Neugeborene in 1 Krankenhaus in Taiwan	k. A.
Kanegae 2016	Neugeborene in 2 Krankenhäusern in São Paulo und Pará, Brasilien	k. A.
de Felipe 2016	Neugeborene in 3 Krankenhäusern in Sevilla, Spanien	k. A.
a: Ergebnisse werden nicht dargestellt. k. A.: keine Angabe		

Tabelle 25: Charakterisierung der Studienpopulationen – Studien zur Testgüte

Studie Gruppe	N	Alter [Lebensstunden] MW (SD)	Geschlecht w / m	Anteil Frühgeborener
Kwan 2015	6100	k. A. ^a	k. A.	k. A.
Vogel 2014 ^b	485 912	k. A. ^c	k. A.	k. A.
Kwan 2013	993 724	k. A. ^d	k. A.	k. A. ^e
Verbsky 2012 ^b	207 696	k. A.	k. A.	9,13 %
Routes 2009 ^b	71 000	k. A. ^c	k. A.	9,3 %
Chien 2015	106 391	k. A. ^c	k. A.	7,9 %
Kanegae 2016	8682	k. A. ^c	k. A.	k. A. ^f
de Felipe 2016	3810 ^g	k. A.	k. A. ^h	k. A. ^h
<p>a: Es wurden Kinder gescreent, die im Alter von 24 Lebensstunden im Krankenhaus waren. b: Ergebnisse werden nicht dargestellt. c: SCID-Screening im Rahmen des Neugeborenen Screenings. d: Probenentnahme im Rahmen des Neugeborenen Screenings; Testung auf SCID nach 20 Tagen (Mittelwert). e: 9 % der Studienpopulation waren zur Zeit des Tests in der Neugeborenen-Intensivstation (NICU). f: Angabe nur für einen Teil der Teilpopulation aus Pará (28 % der Gesamtpopulation; N = 2416); Anteil Frühgeborener: 9,9 % (n = 239). g: Es werden nur die um die Daten der Pilotstudie [33] bereinigten Daten berichtet (3810 von 5160). h: Die Angaben aus der Publikation beziehen sich auf die gesamte Studienpopulation von 5160 Neugeborenen und können der hier berichteten Teilpopulation (N = 3810) nicht zugeordnet werden: Geschlecht: 51,2 % männlich; Anteil Frühgeborener: 25,4 %.</p> <p>MW: Mittelwert; m: männlich; N: Anzahl randomisierter (bzw. eingeschlossener) Personen; SCID: Severe combined immunodeficiency (schwerer kombinierter Immundefekt); SD: Standardabweichung; w: weiblich</p>				

Tabelle 26: Charakterisierung von Index- und Referenztest

Studie	Indextests	Referenztest
Kwan 2015	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwerte: <ul style="list-style-type: none"> ▫ 1. Trennwert (1. Messung): < 40 TRECs/μL ▫ 2. Trennwert (2. Messung): < 25 TRECs/μL ▪ Kontrolle der Probenqualität: β-Actin (normal / abnormal), wenn ≤ 40 TRECs/μL ▪ Ergebnis 1. Messung: <ul style="list-style-type: none"> ▫ > 40 TRECs/μL: normal ▫ ≤ 40 TRECs/μL: 2. Messung inkl. Messung von β-Actin anhand gleicher Filterkarte ▪ Ergebnis 2. Messung: <ul style="list-style-type: none"> ▫ > 25 TRECs/μL: normal ▫ 0 TRECs/μL, β-Actin normal: testpositiv, sofortiger Referenztest ▫ 0 TRECs/μL, β-Actin abnormal: „inconclusive“; sofortige Wiederholungsmessung^a anhand neuer Filterkarte ▫ 1–25 TRECs/ μL, β-Actin normal: „vermutlich testpositiv“; sofortiger Referenztest ▫ 1–25 TRECs/ μL, β-Actin abnormal: „inconclusive“; sofortige Wiederholungsmessung^a anhand neuer Filterkarte 	<ul style="list-style-type: none"> ▪ klinische Untersuchung ▪ immunologisches Profil
Vogel 2014	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwert^b (jeweils durchschnittlicher Wert aus 3 Messungen [gleiche Filterkarte] bzw. Wert aus 2 der 3 Messungen): <ul style="list-style-type: none"> ▫ 1. Trennwert: ≤ 200 TRECs/μL ▫ 2. Trennwert: 150–200 TRECs/μL ▫ 3. Trennwert: 125–200 TRECs/μL ▪ Kontrolle der Probenqualität: RNase P (normal / abnormal) ▪ generell: wenn RNase P abnormal: Wiederholungsmessung ▪ Ergebnis 1. Messung (jeweils, wenn RNase P normal): <ul style="list-style-type: none"> ▫ entsprechend allen Trennwerten: > 200 TRECs/μL: testnegativ^c ▫ entsprechend 1. Trennwert^a: ≤ 200 TRECs/μL: abnormal ▫ entsprechend 2. Trennwert^a: 150–200 TRECs/μL: Wiederholungsmessung anhand neuer Probe ▫ entsprechend 3. Trennwert^a: 125–200 TRECs/μL: Wiederholungsmessung anhand neuer Probe ▪ Ergebnis 2. Messung (jeweils, wenn RNase P normal): <ul style="list-style-type: none"> ▫ 0 TRECs/μL: testpositiv ▫ entsprechend 1. Trennwert^a: ≤ 200 TRECs/μL: <i>abnormal</i> ▫ entsprechend 2. Trennwert^a: <ul style="list-style-type: none"> - ≥ 150/μL und ≤ 200 TRECs/μL: Wiederholungsmessung - < 150/μL: testpositiv ▫ entsprechend 3. Trennwert^a: <ul style="list-style-type: none"> - ≥ 125 TRECs/μL: Wiederholungsmessung - < 125 TRECs/μL: testpositiv 	<ul style="list-style-type: none"> ▪ Vollblutanalyse (großes Blutbild), Durchflusszytometrie ▪ wenn angemessen, stattdessen Bestimmung der Zellaktivierung mittels Mitogentest, Chromosomenanalyse oder Gentest

(Fortsetzung)

Tabelle 26: Charakterisierung von Index- und Referenztest (Fortsetzung)

Studie	Indextests	Referenztest
Kwan 2013	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut^d ▪ Trennwert: < 40 TRECs/μL ▪ Kontrolle der Probenqualität: β-Actin (normal / abnormal), wenn < 40 TRECs/μL ▪ Ergebnis 1. Messung: <ul style="list-style-type: none"> ▫ > 40 TRECs/μL: normal ▫ ≤ 40 TRECs/μL: 2. Messung inkl. Messung von β-Actin^e ▪ Ergebnis 2. Messung: <ul style="list-style-type: none"> ▫ > 25 TRECs/μL: normal ▫ ≤ 5 TRECs, β-Actin normal: testpositiv, sofortiger Referenztest ▫ ≤ 5 TRECs, β-Actin abnormal; Wiederholungsmessung anhand neuer Probe, ggf. Lymphozytensubgruppenbestimmung <ul style="list-style-type: none"> - Wiederholungsmessung anhand neuer Filterkarte: <ul style="list-style-type: none"> - > 25 TRECs/μL: normal - ≤ 25 TRECS testpositiv ▫ 6–25 TRECs/μL, β-Actin normal: testpositiv^f, sofortiger Referenztest ▫ 6–25 TRECs/μL, β-Actin abnormal: „incomplete^g“ <ul style="list-style-type: none"> - Wiederholungsmessung anhand neuer Filterkarte: <ul style="list-style-type: none"> - > 25 TRECs/μL: normal - ≤ 25 TRECS/μL: testpositiv 	<ul style="list-style-type: none"> ▪ Vollblutanalyse (großes Blutbild) mittels Durchflusszytometrie ▪ Lymphozytensubgruppenbestimmung
Verbsky 2012	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwerte^h <ul style="list-style-type: none"> ▫ 1. Trennwert: < 25 TRECs/μL ▫ 2. Trennwert: < 40 TRECs/μL ▪ Ergebnis 1. Messung entsprechend allen Trennwerten: <ul style="list-style-type: none"> ▫ TRECs/μL ≥ [Trennwert]: normal ▫ TRECs/μL < [Trennwert]^k: abnormal; Messung von β-Actin ▫ TRECs/μL < [Trennwert], β-Actin normal: testpositivⁱ ▫ TRECs/μL < [Trennwert], β-Actin niedrig: „inconclusive“: Wiederholungsmessung anhand neuer Filterkarte 	<ul style="list-style-type: none"> ▪ positive Testergebnisse: Lymphozytensubgruppenanalyse mittels Durchflusszytometrie

(Fortsetzung)

Tabelle 26: Charakterisierung von Index- und Referenztest (Fortsetzung)

Studie	Indextests	Referenztest
Routes 2009	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwert: < 25 TRECs/μL ▪ Kontrolle der Probenqualität: β-Actin (normal / niedrig) ▪ Ergebnis 1. Messung: <ul style="list-style-type: none"> ▫ ≥ 25 TRECs/μL: normal ▫ < 25 TRECs/μL: 2. Messung anhand neuer Probe aus derselben Filterkarte ▪ Ergebnis 2. Messung <ul style="list-style-type: none"> ▫ ≥ 25 TRECs/μL: <i>normal</i> ▫ < 25TRECs, β-Actin niedrig^j: „inconclusive“: Wiederholungsmessung anhand neuer Probe ▫ < 25TRECs, β-Actin normal^j: testpositiv ^k 	<ul style="list-style-type: none"> ▪ positive Testergebnisse: Lymphozytensubgruppenanalyse mittels Durchflusszytometrie
Chien 2015	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwert: < 40 TRECs/μL ▪ Kontrolle der Probenqualität: RNase P (normal / abnormal) ▪ Ergebnis 1. Messung: <ul style="list-style-type: none"> ▫ 0 TRECs (normale RNase P): <i>abnormal</i> ▫ 0–40 TRECs: <i>inconclusive</i> ▪ 2. Messung (neue Filterkarte), falls 1. Messung <i>inconclusive</i> ▪ Ergebnis 2. Messung: <ul style="list-style-type: none"> ▫ 0–40 TRECs: <i>abnormal</i> 	<ul style="list-style-type: none"> ▪ positive Testergebnisse: ▪ Testergebnis „<i>abnormal</i>“: Vollblutanalyse (großes Blutbild mittels Durchflusszytometrie) ▪ Testergebnis <i>abnormal</i> oder <i>inconclusive</i>: Test auf Mikrodeletionssyndrom 22q11.2^l
Kanegae 2016	<ul style="list-style-type: none"> ▪ Marker: TREC ▪ Probe: Filterkartenblut ▪ Trennwert: < 30 TRECs/μL ▪ Ergebnis 1. Messung: <ul style="list-style-type: none"> ▫ ≥ 30 TRECs/μL: normal ▫ < 30 TRECs/μL, β-Actin hoch: 2. Messung anhand neuer Probe aus derselben Filterkarte ▪ Ergebnis 2. Messung <ul style="list-style-type: none"> ▫ ≥ 30 TRECs/μL: normal ▫ < 30TRECs, β-Actin niedrig: „inconclusive“: Wiederholungsmessung anhand neuer Probe ▫ < 30TRECs, β-Actin normal: testpositiv 	<ul style="list-style-type: none"> ▪ positive Testergebnisse: Vollblutanalyse sowie Lymphozytensubgruppenanalyse (T-, B- und NK-Zellen)

(Fortsetzung)

Tabelle 26: Charakterisierung von Index- und Referenztest (Fortsetzung)

Studie	Indextests	Referenztest
de Felipe 2016	<ul style="list-style-type: none"> ▪ Marker: TREC und KREC ▪ Probe: Filterkartenblut^m ▪ Trennwert: <ul style="list-style-type: none"> ▫ TREC: < 6 TRECs/Probe ▫ KREC: < 4 KRECs/Probe ▪ Kontrolle der Probenqualität: β-Actin (normal / abnormal) ▪ Ergebnis 1. Messungⁿ: <ul style="list-style-type: none"> ▫ \geq Trennwert TREC oder KREC: normal ▫ < Trennwert TREC und / oder KREC: inconclusiveⁿ oder abnormal^o ▪ 2. Messung, falls 1. Messung inconclusiveⁿ oder abnormalⁿ: 2. Messung anhand neuer Probe aus derselben Filterkarte^p ▪ Ergebnis 2. Messung: <ul style="list-style-type: none"> ▫ < Trennwert TREC oder KREC: weitere Messung anhand neuer Filterkartenprobe ▪ Ergebnis 3. Messung: <ul style="list-style-type: none"> ▫ < Trennwert TREC oder KREC: abnormal 	<ul style="list-style-type: none"> ▪ Lymphozytensubgruppenanalyse mittels Durchflusszytometrie^q

(Fortsetzung)

Tabelle 26: Charakterisierung von Index- und Referenztest (Fortsetzung)

<p>a: Gilt für Termingeborene (≥ 37 Wochen); Frühgeborene erhalten eine Wiederholungsmessung, sobald das Alter der 37. Gestationswoche erreicht wird.</p> <p>b: Im Rahmen des Studienzeitraums wurde nach 4 Monaten (Januar 2011) eine <i>borderline</i>-Kategorie (TRECs/μL 150–200) eingeführt, die nach weiteren 6 Monaten (Juli 2011) nochmals angepasst wurde (TRECs/μL 125–200).</p> <p>c: gilt für Früh- und Termingeborene</p> <p>d: Die Filterkarten wurden im Mittel 20 Tage nach der Geburt auf SCID getestet; auffällige Ergebnisse wurden in der Regel 3–4 Wochen nach der Geburt mitgeteilt.</p> <p>e: Abweichend bei 13 Neugeborenen, bei denen wegen klinischer Auffälligkeit bereits nach erstem auffälligem Testergebnis der Referenztest durchgeführt wurde.</p> <p>f: gilt für Termingeborene</p> <p>g: gilt für Frühgeborene; keine Angabe zum Gestationsalter als Differenzierungsmerkmal</p> <p>h: Im Rahmen des Studienzeitraums wurde nach 19 Monaten der Trennwert aufgrund der geringen Anzahl identifizierter Kinder mit einem als <i>abnormal</i> klassifizierten Testergebnis angehoben.</p> <p>i: In diesem Fall ist die Bestätigung des Testergebnisses durch den Referenztest vorgesehen; einige Kliniker haben in diesem Fall jedoch eine zweite Messung angefordert. Konkrete Angaben zur Handhabung liegen nicht vor.</p> <p>j: Gilt für Termingeborene; bei Frühgeborenen < 37. Gestationswoche erfolgten bei einem als <i>abnormal</i> oder <i>inconclusive</i> eingestuften Ergebnis Wiederholungsmessungen bis zum Erreichen eines normalen Testergebnisses oder des Alters entsprechend der 37. Gestationswoche mit gegebenenfalls Reklassifizierung als testpositiv bei Testergebnis <i>abnormal</i>.</p> <p>k: In diesem Fall kann entweder die Bestätigung des Testergebnisses durch den Referenztest erfolgen oder aber eine zweite Messung angefordert werden. Konkrete Angaben zur Handhabung liegen nicht vor.</p> <p>l: Sogenannter TUPLE1-Test; möglich ist der Nachweis beispielsweise des DiGeorge-Syndroms, aber auch anderer Krankheiten, die nicht unter SCID fallen.</p> <p>m: Aus den Filterkarten werden für den Test 2 Proben pro Kind ausgestant.</p> <p>n: Gilt für Termingeborene; bei Frühgeborenen < 37. Gestationswoche alle 2 Wochen Wiederholungsmessungen bis zum Erreichen eines normalen Testergebnisses oder Alter entsprechend der 37. Gestationswoche oder Geburtsgewicht ≥ 2500 g.</p> <p>o: Keine Definition von <i>inconclusive</i> und <i>abnormal</i> in der Hauptpublikation. In der zugehörigen Pilotstudie [33] wurde bei auffälligem Wert der Kontrolle (Ergebnis: abnormal) das Messergebnis als <i>inconclusive</i> bezeichnet.</p> <p>p: Für den Fall, dass aus derselben Filterkarte keine neue Probe gestant werden kann, kann eine neue Filterkartenprobe angefordert werden.</p> <p>q: Der Referenztest ist in den Methoden nicht genannt; die Angaben stammen aus der Beschreibung der positiven Testergebnisse. In der zugehörigen Pilotstudie [33] wurden als Referenztests klinische Untersuchung und Durchflusszytometrie angegeben.</p> <p>KREC: Kappa-deleting Recombination Excision Circles; RNase P: Ribonuklease P; SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); TREC: T-Cell Receptor Excision Circles</p>

A3.6.2 Einschätzung des Verzerrungspotenzials

Die Bestimmung des Verzerrungspotenzials und der Übertragbarkeit der Ergebnisse der Primärstudien auf die Fragestellung des Berichts erfolgt auf Basis des Instruments QUADAS 2 [47] (siehe Abschnitt A2.1.6.2).

A3.6.2.1 Verzerrungspotenzial nach QUADAS 2

Im Folgenden sind die Ergebnisse der Bewertung des Verzerrungspotenzials von 5 Studien dargestellt, deren Ergebnisse für den Bericht verwertet wurden. Für die Studie Chien 2015 [30] wurde zur Beurteilung des Verzerrungspotenzials eine Autorenanfrage gestellt; die Antwort ist in die Beurteilung eingeflossen. Die 5 Studien weisen ein hohes Ver-

zerrungspotenzial auf. Ausschlaggebend für diese Bewertung sind die jeweils unvollständige Darstellung des Patientenflusses sowie die Anwendung des Referenzstandards in Kenntnis des Testergebnisses. Zwar ist nachvollziehbar, dass jeweils nur die auffällig getesteten Kinder mittels Referenztest nachuntersucht wurden. Das ändert aber nichts daran, dass die zwangsläufige Kenntnis des positiven Testergebnisses die Interpretation des Ergebnisses des Referenzstandards beeinflusst haben kann.

Tabelle 27: Verzerrungspotenzial nach QUADAS 2

Studie	Patientenselektion (Domäne 1)	Indextest I (Domäne 2)	Referenzstandard (Domäne 3)	Patientenfluss und zeitl. Ablauf (Domäne 4)	Zusammenfassende Einschätzung
Kwan 2015	niedrig	niedrig	unklar ^a	hoch ^b	hoch
Vogel 2014	entfällt ^c				
Kwan 2013	niedrig	niedrig	unklar ^a	hoch ^d	hoch
Verbsky 2012	entfällt ^c				
Routes 2009	entfällt ^c				
Chien 2015	niedrig	niedrig	unklar ^a	hoch ^e	hoch
Kanegae 2016	niedrig	niedrig	unklar ^a	unklar ^f	hoch
de Felipe 2016	niedrig	niedrig	unklar ^a	unklar ^g	hoch

a: Die Anwendung des Referenzstandards erfolgte in Kenntnis des Testergebnisses.
b: Der Patientenfluss ist nicht vollständig abgebildet: Das Flussdiagramm berichtet die Angaben von 3733 von 6100 Neugeborenen; weitere Angaben im Text decken die Kategorien der Testergebnisse nicht ab; insbesondere fehlen Angaben zur Kategorie „presumptive positive“-Testergebnisse (1–25 TREC).
c: Die Ergebnisse werden nicht dargestellt, daher entfällt die Bewertung des Verzerrungspotenzials der Ergebnisse.
d: Zwischen dem Test auf SCID und dem Referenztest lag eine durchschnittliche Zeitspanne von 3 bis 4 Wochen. Die Autoren verweisen darauf, dass in der früheren postpartalen Phase die Anzahl von TRECs steigt, und begründen das Zeitintervall mit der Möglichkeit der Auflösung einer ursprünglichen T-Zell-Lymphopenie bei einem geringen Risiko für Neugeborene mit SCID. Ferner wurden aufgrund von Versterben nicht alle positiven (n = 11) oder als *inconclusive* kategorisierten Testergebnisse (n = 3) mittels Referenztest verifiziert; ein weiteres positives Testergebnis wurde ebenfalls nicht überprüft (Lost to Follow-up).
e: Von 90 eingeschlossenen Neugeborenen fehlt die Angabe des Testergebnisses (keine Angabe der Gründe); bei weiteren 60 Frühgeborenen wurde das ursprünglich als *inconclusive* eingeordnete Testergebnis nicht endgültig interpretiert, weil das Alter der 37. Gestationswoche nicht erreicht war (Angabe der Autoren).
f: Es werden nur die als *abnormal* eingeordneten Ergebnisse berichtet; es fehlen Angaben zu *inconclusive* und *normal*-Ergebnissen.
g: Es können nur die als „abnormal“ eingeordneten Ergebnisse der berichteten Teilpopulation zugeordnet werden; für diese Gruppe fehlen Angaben zu *inconclusive* und *normal*-Ergebnissen.
SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)

A3.6.2.2 Bedenken bezüglich der Übertragbarkeit nach QUADAS 2

Die Übertragbarkeit der 5 Studien, deren Ergebnisse im Bericht dargestellt werden, ist in der folgenden Tabelle dargestellt. In den Studien mit starken Bedenken hinsichtlich der Übertragbarkeit ist die – im Vergleich zu dem im vorliegenden Versorgungskontext bekannte – hohe Prävalenz für SCID in den Studienpopulationen maßgeblich. Hinzu kommt, dass die Anwendung und Interpretation des Referenzstandards (Durchflusszytometrie) abhängig von lokalen Erfahrungen und schon aus diesem Grund unklar ist, ob ein im deutschen Versorgungskontext eingesetzter entsprechender Referenztest vergleichbar wäre.

Tabelle 28: Bedenken bezüglich der Übertragbarkeit nach QUADAS 2

Studie	Patientenselektion (Domäne 1)	Indextest I (Domäne 2)	Referenzstandard (Domäne 3)	Zusammenfassende Einschätzung
Kwan 2015	stark ^a	gering	unklar ^b	stark
Vogel 2014	entfällt ^c			
Kwan 2013	gering	gering	unklar ^b	gering
Verbsky 2012	entfällt ^c			
Routes 2009	entfällt ^c			
Chien 2015	stark ^d	gering	unklar ^b	stark
Kanegae 2016	gering	gering	unklar ^b	gering
de Felipe 2016	gering	gering	unklar ^b	gering
a: Die Inzidenz von SCID in der untersuchten Population ist sehr hoch und wird mit etwa 1:1580 als 30-mal höher als in anderen Populationen beschrieben. b: Die Anwendung und Interpretation des Referenzstandards (Durchflusszytometrie) ist abhängig von lokalen Erfahrungen. c: Die Ergebnisse werden nicht dargestellt, daher entfällt die Bewertung der Übertragbarkeit der Ergebnisse. d: Die Inzidenz wird mit über 1:11 821 als wesentlich höher als in Deutschland geschätzt angesehen. SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt)				

A3.7 Ergebnisse zu den Zielgrößen

Zielgrößen für die Nutzenbewertung des SCID-Screenings waren Maße der diagnostischen Güte in Hinblick auf die Entdeckung von Neugeborenen mit SCID.

Die Ergebnisse zur Testgüte aus 5 Studien sind in der folgenden Tabelle beschrieben. Bei **Chien 2015** ist die Antwort auf die Autorenanfrage hinsichtlich in der Publikation enthaltener Angaben zur Anzahl getesteter Neugeborener und zur Anzahl von Testergebnissen einge-

gangen. Bei **de Felipe 2016** ist die Autorenantwort hinsichtlich der um die Daten aus der Pilotstudie bereinigten Daten eingegangen (siehe Abschnitt A3.1.2.4).

Die Ergebnisse der 5 ausgewerteten Studien sind heterogen. Für den deutschen Versorgungskontext scheinen die Ergebnisse aus der Studie Kwan 2013 am aussagekräftigsten.

Die Ergebnisse der übrigen 3 eingeschlossenen Studien waren nicht interpretierbar und werden nicht dargestellt.

In der Studie **Vogel 2014** [28] wurden während der Studienlaufzeit unter dem ursprünglichen Grenzwert von 200 TREC's 2 sogenannte *borderline*-Kategorien eingeführt. Nach welchem Grenzwert die Ergebnisse ausgewertet wurden, bleibt unklar. Bei Zugrundelegung eines Grenzwertes von 200 TREC's (1745 positive Testergebnisse, abzüglich der Ergebnisse von 541 Frühgeborenen) bleibt die Datenanalyse unvollständig; mit 531 Verifizierungen wurde noch nicht einmal die Hälfte der positiven Testergebnisse mittels Referenztest überprüft. Eine Autorenanfrage wurde nicht gestellt.

Ähnliches gilt für die Publikationen **Verbsky 2012** [29] und **Routes 2009** [17]. Die Publikation Verbsky 2012 enthält die von Routes berichteten Daten. Innerhalb des von Verbsky berichteten Studienzeitraums wurde der Trennwert angepasst; eine Zuordnung der berichteten Ergebnisse zu den jeweils verwendeten Trennwerten ist nicht möglich. Wegen fehlender Angaben zum Patientenfluss ist unklar, wie viele der positiven Testergebnisse mittels Referenztest überprüft wurden. Eine Autorenanfrage mit der Bitte um eine Ergänzung der publizierten Daten blieb inhaltlich unbeantwortet.

Die unvollständige Anwendung des Referenztests ergibt sich für einen Teil der von Verbsky 2012 berichteten Daten aus der Studie von Routes 2008. In der Gruppe der Früh- und Termingeborenen haben nur 11 von 42 positiv getesteten Neugeborenen den Referenztest erhalten.

Zusammenfassend spricht die dürftige Datenlage für einen in Deutschland zu erwartenden PPV, der in etwa zwischen 4 % und 13 % liegen dürfte. Basis hierfür sind die Daten aus der mit Abstand größten Studie Kwan 2013, die insbesondere angesichts der Inzidenz dem deutschen Versorgungskontext entsprechen.

Tabelle 29: Ergebnisse der Studien zur diagnostischen Güte des Screeningtests auf SCID

Studie	n	Index-test	Trennwert	Referenztest	RP	FN	FP	RN	PPV in % [95 %-KI]
Kwan 2015	6100	TREC	> 40 TRECs (1. Messung) > 25 TRECs (2. Messung)	klinische Untersuchung, immunologisches Profil	4 ^a	-	0	-	100 [39,8; 100] ^b
Vogel 2014	485 912	TREC	<ul style="list-style-type: none"> ▪ 1. Trennwert: durchschnittlicher Wert aus 3 Messungen beziehungsweise Wert aus 2 der 3 Messungen ≤ 200 TRECs/μL ▪ 2. Trennwert: 150–200 TRECs/μL ▪ 3. Trennwert: 125–200 TRECs/μL 	Vollblutanalyse (großes Blutbild), Durchflusszytometrie; wenn angemessen, stattdessen Bestimmung der Zellaktivierung mittels Mitogentest, Chromosomenanalyse oder Gentest	-	-	-	-	
Kwan 2013	993 724	TREC	> 40 TRECs (1. Messung) > 25 TRECs (2. Messung) ^c	Vollblutanalyse (großes Blutbild) mittels Durchflusszytometrie; Lymphozyten-subgruppenbestimmung	12 ^d	-	149	-	7,5 [3,9; 12,7] ^b
Chien 2015	106 391	TREC	≥ 40 TRECs/μL ^e	Vollblutanalyse (großes Blutbild mittels Durchflusszytometrie), Test auf Mikrodeletions-syndrom 22q11.2	2 ^f	-	22	-	8,3 [1,0; 27,0] ^b
Routes 2009	71 000	TREC	> 25 TRECs	Lymphozyten-subgruppenanalyse mittels Durchflusszytometrie	-	-	-	-	

(Fortsetzung)

Tabelle 29: Ergebnisse der Studien zur diagnostischen Güte des Screeningtests auf SCID (Fortsetzung)

Studie	n	Index-test	Trennwert	Referenztest	RP	FN	FP	RN	PPV in % [95 %-KI]
Verbsky 2012	207 696	TREC	<ul style="list-style-type: none"> ▪ 1. Trennwert: < 25 TRECs/μL ▪ 2. Trennwert: < 40 TRECs/μL 	Lymphozyten-subgruppenanalyse mittels Durchflusszytometrie	-	-	-	-	
Kanegae 2016	8682	TREC	<ul style="list-style-type: none"> ▪ ≥ 30 TRECs (1. Messung) ▪ ≥ 30 TRECs (2. Messung) 	klinische Untersuchung, immunologisches Profil	0 ^g	-	3 ^g	-	0 [0; 70,8] ^b
de Felipe 2016	3810	TREC und KREC	<ul style="list-style-type: none"> ▪ < 6 TRECs und / oder ▪ < 4 KRECs 	Lymphozyten-subgruppenanalyse mittels Durchflusszytometrie	0	-	4	-	0 [0; 60,2] ^b

a: RP mit der Diagnose homozygote Mutation für SCID-A
b: eigene Berechnung
c: 15 positiv getestete Neugeborene gingen nicht in die Auswertung ein (14 Todesfälle, 1 Lost to Follow-up)
d: RP mit den Diagnosen: IL2RG (n = 4); JAK3 (n = 1); IL7R (n = 3); RAG1 (n = 4). Die weiteren berichteten Diagnosen fallen nicht unter die Definition eines SCID im Sinne des vorliegenden Berichts (RMRP I [n = 1], 22q11 Deletion [n = 1]; 1 Patient verstarb mit unbekannter genetischer Diagnose [Lymphocyt-Phenotyp: T-B+NK+]).
e: testpositiv („abnormal“) bereits nach 1. Messung: n = 5; testpositiv („abnormal“) nach 2 Messungen: n = 19
f: RP mit den Diagnosen: IL2RG, RAG1. Die weiteren berichteten Diagnosen fallen nicht unter die Definition eines SCID im Sinne des vorliegenden Berichts (SCID Variants; n = 2).
g: Lost to Follow-up: n = 1
FN: falsch-negativ; FP: falsch-positiv; KI: Konfidenzintervall; PPV: Positive predictive Value (positiver prädiktiver Wert); RN: richtig-negativ; RP: richtig-positiv, SCID: Severe combined Immunodeficiency (schwerer kombinierter Immundefekt); TREC: T-Cell Receptor Excision Circles

A3.7.1 Meta-Analysen

Es wurde keine Meta-Analyse durchgeführt.

A3.7.2 Sensitivitätsanalysen

Es wurde keine Sensitivitätsanalyse durchgeführt.

A3.7.3 Subgruppenmerkmale und andere Effektmodifikatoren

Es wurde keine Subgruppenanalyse durchgeführt.

A4 Kommentare

Nachfolgend werden die Ergebnisse der vorliegenden Nutzenbewertung kommentiert. Sofern thematisch zutreffend werden dabei Aspekte aus der Anhörung zum Vorbericht gewürdigt. Eine Liste aller wesentlichen Aspekte aus der Anhörung zum Vorbericht findet sich in Abschnitt A4.3. Außerdem werden in diesem Abschnitt die Aspekte gewürdigt, die in den Abschnitten A4.1 und A4.2 noch nicht adressiert wurden.

A4.1 Bericht im Vergleich zu anderen systematischen Übersichten

Im Rahmen der systematischen Literaturrecherche wurden 2 systematische Übersichten identifiziert (**Lipstein 2010** [14] und **van der Spek 2015** [64]). Darüber hinaus wurde 1 weitere systematische Übersichtsarbeit für das UK National Screening Committee identifiziert (**Bazian 2012**) [4], die auf der Arbeit von Lipstein aufbaut und diese aktualisiert.

In **Lipstein 2010** werden, wie im vorliegenden Bericht, einzelne Bausteine der Screeningkette untersucht. Für die Bewertung **frühere versus spätere Intervention** haben die Autoren 4 Studien eingeschlossen, darunter 3, die im Rahmen der vorliegenden Erstrecherche identifiziert, jedoch nicht in die Nutzenbewertung eingeschlossen wurden [65-67]. Die von Lipstein eingeschlossene Studie Myers 2002 [25] wurde auch in die vorliegende Bewertung eingeschlossen. In der Aktualisierung von **Bazian 2012** sind zum Vergleich frühere versus spätere Therapie 5 weitere Studien eingeschlossen. Es handelt sich neben der für das Fazit des vorliegenden Berichts entscheidenden Studie Brown 2011 [24] um 4 weitere Studien, die vorliegend ausgeschlossen wurden, überwiegend wegen fehlender Angaben dazu, ob die früh behandelten Patienten tatsächlich durch ein Screening oder eine vergleichbare Diagnose-situation und nicht etwa wegen besonders schwerer SCID-Symptome gefunden wurden [68-71].

Für den Vergleich frühere versus spätere Therapie kommen Lipstein et al. zu einer positiven Bewertung einer früheren Stammzelltherapie, ohne dieses Ergebnis jedoch zu spezifizieren. Bazian stützt dieses Ergebnis mit dem Zusatz, dass der Vorteil in einer früheren kurativen Behandlung bestehe, wenn bis dahin keine Infektionen aufgetreten seien.

Für die Bewertung der **diagnostischen Güte** stützt sich **Lipstein 2010** auf 3 Fall-Kontroll-Studien und 1 Kohortenstudie, die die Einschlusskriterien der vorliegenden Nutzenbewertung nicht erfüllten beziehungsweise bereits auf Titel- / Abstractebene im Referenzencheck als nicht relevant ausgeschlossen wurden. Für die diagnostische Güte haben sie eigene Berechnungen angestellt und präsentieren Daten zur Sensitivität und Spezifität. Diese sind nicht überprüfbar, weil sie zum Teil auf unveröffentlichten Daten (schriftlichen Mitteilungen) beruhen.

Bazian 2012 beschreibt darüber hinaus noch die hier nicht dargestellten Ergebnisse der eingeschlossenen Studien Verbsky 2012 [29] und Routes 2009 [17] sowie einiger weiterer

Studien, die vorliegend im Wesentlichen als Validierungsstudien ausgeschlossen wurden [15,16,72,73].

Die Übersichtsarbeit **van der Spek 2015** fokussiert auf die Bewertung der diagnostischen Güte eines Tests auf SCID. Diese Arbeit stellt, wie der vorliegende Bericht, nur Werte zum PPV dar, ohne die Daten in einer Meta-Analyse zusammenzufassen. Es wird auf die große Spannweite der publizierten Ergebnisse verwiesen.

Im Vergleich zu den 3 systematischen Übersichten wurde für die vorliegende Nutzenbewertung ein kleinerer Studienpool herangezogen. Gleichwohl deuten die Ergebnisse in die gleiche Richtung.

A4.2 Kritische Reflexion des Vorgehens

Die Anforderungen an das Evidenzlevel wurden stark reduziert, und zwar aufgrund der Seltenheit und der Schwere der Erkrankung sowie wegen der Möglichkeit, Studien zu identifizieren, die einen dramatischen Effekt zeigen. Dennoch liegt dem vorliegenden Bericht im Vergleich zu den anderen systematischen Übersichten ein kleinerer Studienpool zugrunde. Das liegt an der engeren Formulierung der Einschlusskriterien.

So wird vorliegend für den Einschluss als **Studie zum Vergleich früherer versus späteren Therapiebeginn** verlangt, dass in die Gruppe des früheren Therapiebeginns Patienten eingeschlossen wurden, die ähnlich wie in einer Screeningsituation identifiziert wurden. Um dies beurteilen zu können, waren konkrete Angaben zur Diagnose erforderlich, insbesondere zum Diagnosezeitpunkt oder zum Zustand des Neugeborenen (symptomatisch / asymptomatisch). Auszuschließen waren Studien, die zwar einen Vergleich zwischen verschiedenen Therapiezeitpunkten zuließen, aber keine entsprechenden Informationen zur Diagnose (Zeitpunkt und Anlass) boten. Dieses Vorgehen rechtfertigt sich aus der Überlegung heraus, dass eine Aussage für eine Kombination aus Screening und Therapie zu treffen war. Entscheidend für Screeningmaßnahmen ist, dass sie Betroffene bereits im asymptomatischen Zustand identifizieren können, um sie entsprechend früher einer Therapie zuführen zu können.

Die Anzahl der **Studien zur diagnostischen Güte** wurde insbesondere durch die Vorgabe bestimmt, dass die laboranalytische Methodik und der Grenzwert zur Unterscheidung positiver und negativer Ergebnisse prospektiv festgelegt und an einer unabhängigen Stichprobe entwickelt worden sein müssen. Aufgrund der Seltenheit der Erkrankung ist einerseits verständlich, wenn Studien dieses Kriterium nicht erfüllen, andererseits sind auf solcher Basis jedoch keine verlässlichen Aussagen zur diagnostischen Güte möglich.

Hervorzuheben als in die Übersichtsarbeit van der Spek 2015, nicht aber in die vorliegende Nutzenbewertung eingeschlossene Studie ist Kwan 2014 [3]. In dieser Studie sind Daten aus 11 US-amerikanischen Staaten zusammengetragen, in denen das Screening auf SCID implementiert ist. So kommen die Autoren auf eine Zahl von 3 030 083 gescreenten

Neugeborenen. Daten von 4 Staaten wurden auch vorliegend aus den entsprechenden Primärstudien eingeschlossen [17,23,27-29]. Allerdings waren die Daten aus 3 dieser Publikationen aufgrund der unvollständigen Anwendung des Referenztestes nicht interpretierbar. Hinweise auf diese Defizite fehlen in der Publikation Kwan 2014. In der Zusammenschau mit den Unterschieden der 11 angewandten Screeningalgorithmen sind die präsentierten Daten nicht verlässlich. Im Übrigen geben diese eine ähnliche Spannbreite der diagnostischen Güte wieder wie der vorliegende Bericht.

Daten zur Sensitivität und Spezifität sind auch aus den dort dargestellten Ergebnissen nicht ableitbar. Zwar wurden bei Kwan 2014 keine falsch-negativen Befunde berichtet, aber die unsystematische Erfassung möglicherweise übersehener SCID-Fälle in dieser Studie erlaubt es nicht, hieraus auf eine 100%ige Sensitivität zu schließen.

A4.3 Würdigung der Stellungnahmen

Insgesamt wurden 6 Stellungnahmen zum Vorbericht frist- und formgerecht eingereicht.

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente werden im Folgenden diskutiert. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, z. B. zu rechtlichen Vorgaben für das Institut, angesprochen. Auf solche Punkte wird im Rahmen dieser projektspezifischen Würdigung der Anhörung nicht weiter eingegangen.

Die in der Stellungnahme angesprochenen Aspekte werden in den nachfolgenden Abschnitten A4.3.1 bis A4.3.3 gewürdigt.

Die Zusammenfassung aller Änderungen des Abschlussberichts gegenüber dem Vorbericht, die sich u. a. durch die Anhörung zum Vorbericht ergeben haben, ist in Abschnitt A1.2 bzw. A2.2 dargestellt.

A4.3.1 Würdigung des zusätzlichen Arguments: Gentherapie bei ADA-SCID

Zwei Stellungnahmen beziehen sich auf die Gentherapie als Behandlungsoption bei ADA-Mangel. Beide verweisen auf die Zulassung eines Gentherapeutikums durch die Europäische Zulassungsbehörde (EMA). Ein Stellungnehmender hebt in diesem Zusammenhang die Notwendigkeit hervor, auch ADA-SCID kurz nach der Geburt zu erkennen.

In Kapitel 1 (Hintergrund) ist der ADA-Mangel ausdrücklich als Zielerkrankung eines SCID-Screenings genannt, Gleiches gilt für die Gentherapie als Therapieoption, sodass diese Aspekte bereits berücksichtigt sind. Der zwischenzeitlich erteilten Zulassung durch die EMA wurde durch Einfügen eines entsprechenden Verweises in Kapitel 1 (Hintergrund) Rechnung getragen. Darüber hinaus ist keine Änderung erforderlich.

A4.3.2 Würdigung der zusätzlichen Argumente zum relevanten Studienpool

A4.3.2.1 Berücksichtigung von Kwan 2014 und Pai 2014

Die Stellungnehmenden einer Stellungnahme halten den Ausschluss der Studie Kwan 2014 [3] für ungerechtfertigt. Sie plädieren für die Beachtung der bisher größten Studie zum SCID-Screening.

Des Weiteren zitieren sie die ebenfalls ausgeschlossene Studie Pai 2014 [74] und verweisen auf die Möglichkeit, die beiden Studien als einen historischen Vergleich einer Strategie mit SCID-Screening versus kein Screening zu verwenden.

Die Gründe für den Ausschluss als Studie zur diagnostischen Güte sind in Abschnitt A4.2 dargelegt. Gleichzeitig wurde darauf verwiesen, dass die in der Studie Kwan 2014 wiedergegebene Spannweite der diagnostischen Güte der Spannweite der diagnostischen Güte der eingeschlossenen und ausgewerteten Studien des Berichts entspricht.

Die Studie Pai 2014 vergleicht 2 Gruppen von Kindern mit unterschiedlichem Alter zum Therapiebeginn (kleiner versus größer 3,5 Monate). Sie wurde ausgeschlossen, weil sie mangels Angabe zur Art der Diagnose (durch Screening beziehungsweise eine vergleichbare Situation oder aber durch klinische Präsentation) keine Daten zum Vergleich frühere versus spätere Therapie im Sinne des Berichts liefert.

Die Verwendung von Pai 2014 als historische Kontrolle in Zusammenschau mit der Studie Kwan 2014 für einen Vergleich Screening versus kein Screening ist ebenfalls nicht aussagekräftig: Das Institut verwendet zur Nutzenbewertung von Interventionen primär direkt vergleichende Studien. Aussagen für die Nutzenbewertung werden daher i. d. R. nur aus Ergebnissen direkt vergleichender Studien abgeleitet [40]. Entsprechend war ein im Rahmen des Berichts selbst durchzuführender Vergleich in der Berichtsmethodik nicht vorgesehen. Gründe, von der geplanten Berichtsmethodik abzuweichen, sind nicht ersichtlich.

Vorliegend kommt noch hinzu, dass Pai 2014 die 5-Jahres-Mortalität berichtet, während Kwan 2014 die Mortalität für ein Alter der Kinder „to Age ≥ 11 Month“ angibt. Ferner gibt es eine Überschneidung der Berichtszeiträume von 2 Jahren, wobei die Studiendauer bei Kwan 2014 bei ca. 5 Jahren liegt. Damit ist nicht auszuschließen, dass es Überschneidungen beider Studienpopulationen gibt.

Studien, in denen ein historischer Vergleich durchgeführt wird, indem neuere Daten denen aus einem früheren Zeitraum gegenübergestellt werden, hätten prinzipiell eingeschlossen werden können.

Es ist keine Änderung erforderlich.

A4.3.2.2 Berücksichtigung von Grazioli 2014, la Marca 2014 und Speckmann 2012

Unter Hervorhebung der Notwendigkeit eines Screenings auf SCID unter Verwendung eines kombinierten Tests auf TRECs und KRECs wird in einer Stellungnahme um die Bewertung eines solchen Screenings gebeten. Die Stellungnehmenden verweisen exemplarisch auf 3 Publikationen, die die Verwendung eines entsprechenden kombinierten Tests beschreiben. In einer weiteren Stellungnahme wird ebenfalls auf die Notwendigkeit des Screenings auf B- und T-Zellen-Defekte hingewiesen.

Die Bewertung von Screenings auch mit kombiniertem Test auf TREC und KREC war vorgesehen. Die von den Stellungnehmenden genannten Studien wurden in der systematischen Recherche identifiziert, waren jedoch als nicht relevant für die vorliegende Frage einzustufen: Grazioli 2014 [75] und Speckmann 2012 [76] aufgrund des Studiendesigns als Fallberichte; die Studie la Marca 2014 [77], weil sie von Tests auf PNP-Mangel mithilfe unterschiedliche Biomarker berichtet und nur unter anderem auch TREC und KREC an Filterkartenblut von Neugeborenen sowie anderen Personen mit bekannten Krankheitsstatus. Die Daten der einzigen identifizierten entsprechenden Publikation de Felipe 2016 [32] wurden nach zwischenzeitlichem Eingang der Autorenantwort in den Abschlussbericht eingefügt. Es ist keine weitere Änderung erforderlich.

A4.3.3 Würdigung des zusätzlichen Arguments: Ergebnis der Nutzenbewertung

Eine Stellungnehmende kritisiert die Ableitung eines Nutzens für ein SCID-Screening aufgrund der Datenlage. Sie hebt insbesondere hervor, dass Angaben über die Anzahl von Neugeborenen fehlen, die in dem Screening unentdeckt bleiben.

Wegen der durchgängig fehlenden Nachbeobachtung negativ getesteter Kinder wurden die Studien im VOPT-Design ausgewertet, wie in der berichtsspezifischen Methodik vorgesehen. Es werden lediglich Daten zum PPV berichtet. Die Tatsache fehlender Daten zur Anzahl durch das Screening nicht identifizierter Kinder ist durchgängig adressiert, so im Fazit (Kernaussage und Kapitel 6), im Ergebnisteil (Abschnitt 4.2.3) sowie in der Einordnung des Arbeitsergebnisses (Kapitel 5). Unabhängig von der Anzahl nichtidentifizierter Kinder mit SCID profitieren jedoch diejenigen Neugeborenen mit SCID von dem Screening, die früher identifiziert werden. Dagegen haben Neugeborene mit falsch-negativem Ergebnis aus einem SCID-Screening keinen Schaden zu erwarten im Vergleich zu einer Diagnosestrategie ohne SCID-Screening, weil ein negatives Testergebnis Neugeborene nicht von den übrigen Untersuchungen nach bisherigem diagnostischem Standard ausschließt (siehe Kapitel 5). Entsprechend der für den Bericht festgelegten Methodik war aus den Ergebnissen der eingeschlossenen Studien ein Anhaltspunkt für einen Nutzen des Screenings mit den genannten Einschränkungen abzuleiten. Es ist keine Änderung erforderlich.

A5 Literatur

1. Statistisches Bundesamt. Diagnosedaten der Krankenhäuser ab 2000 (Eckdaten der vollstationären Patienten und Patientinnen); Gliederungsmerkmale: Jahre, Behandlungs-/Wohnort, ICD10; D81.0 Schwerer kombinierter Immundefekt (SCID) mit retikulärer Dysgenese [online]. 10.02.2015 [Zugriff: 19.10.2015]. URL: <https://www.gbe-bund.de>.
2. Verbsky J, Thakar M, Routes J. The Wisconsin approach to newborn screening for severe combined immunodeficiency. *J Allergy Clin Immunol* 2012; 129(3): 622-627.
3. Kwan A, Abraham RS, Currier R, Brower A, Andruszewski K, Abbott JK et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. *JAMA* 2014; 312(7): 729-738.
4. Bazian. Screening for severe combined immunodeficiency: external review against programme appraisal criteria for the UK National Screening Committee (UK NSC); version 3 [online]. 06.2012 [Zugriff: 20.10.2015]. URL: http://legacy.screening.nhs.uk/policydb_download.php?doc=241.
5. Al-Herz W, Bousfiha A, Casanova JL, Chatila T, Conley ME, Cunningham-Rundles C et al. Primary immunodeficiency diseases: an update on the classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency. *Front Immunol* 2014; 5: 162.
6. 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: 17.09.2015]. URL: <http://www.dimdi.de/static/de/klassi/icd-10-gm/kodesuche/onlinefassungen/htmlgm2013/index.htm>.
7. Schwarz K, Fugmann SD, Pannicke U, Bartram CR. Angeborene Immundefekte. In: Ganten D, Bartram CR (Ed). *Immunsystem und Infektiologie*. Berlin: Springer; 1999. S. 40-83. (Handbuch der Molekularen Medizin; Band 4).
8. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über Schutzimpfungen nach § 20d Abs. 1 SGB V (Schutzimpfungs-Richtlinie/SI-RL) [online]. 06.02.2016 [Zugriff: 11.10.2016]. URL: https://www.g-ba.de/downloads/62-492-1118/SI-RL_2015-11-27_iK-2016-02-06.pdf.
9. Albert M, Borte S, Wahn V. Konzept für ein bundesweites Modellprojekt zum Neugeborenen-Screening auf angeborene Immundefekte. *KinderImmunologische Mitteilungen* 2013; 2(2): 12-18.
10. Mannhardt-Laakmann W, Habermehl P, Zepp F. T-zelluläre und kombinierte T- und B-Zell-Immundefekte. *Kinder- und Jugendmedizin* 2008; 8(3): 137-149.
11. Buelow BJ, Routes JM, Verbsky JW. Newborn screening for SCID: where are we now? *Expert Rev Clin Immunol* 2014; 10(12): 1649-1657.

12. Farmand S, Baumann U, Von Bernuth H, Borte M, Förster-Waldl E, Franke K et al. Diagnostik von primären Immundefekten; interdisziplinäre AWMF-Leitlinie (S2k). *Kinder- und Jugendmedizin* 2012; 12(1): 31-40.
13. European Society for Immunodeficiencies. ESID Registry: working definitions for clinical diagnosis of PID [online]. 11.05.2015 [Zugriff: 17.09.2015]. URL: http://esid.org/content/download/13053/372959/file/ESIDRegistry_ClinicalCriteria.pdf.
14. Lipstein EA, Vorono S, Browning MF, Green NS, Kemper AR, Knapp AA et al. Systematic evidence review of newborn screening and treatment of severe combined immunodeficiency. *Pediatrics* 2010; 125(5): e1226-e1235.
15. Baker MW, Grossman WJ, Laessig RH, Hoffman GL, Brokopp CD, Kurtycz DF et al. Development of a routine newborn screening protocol for severe combined immunodeficiency. *J Allergy Clin Immunol* 2009; 124(3): 522-527.
16. Chan K, Puck JM. Development of population-based newborn screening for severe combined immunodeficiency. *J Allergy Clin Immunol* 2005; 115(2): 391-398.
17. Routes JM, Grossman WJ, Verbsky J, Laessig RH, Hoffman GL, Brokopp CD et al. Statewide newborn screening for severe T-cell lymphopenia. *JAMA* 2009; 302(22): 2465-2470.
18. Gerstel-Thompson JL, Wilkey JF, Baptiste JC, Navas JS, Pai SY, Pass KA et al. High-throughput multiplexed T-cell-receptor excision circle quantitative PCR assay with internal controls for detection of severe combined immunodeficiency in population-based newborn screening. *Clin Chem* 2010; 56(9): 1466-1474.
19. Griffith LM, Cowan MJ, Notarangelo LD, Kohn DB, Puck JM, Pai SY et al. Primary Immune Deficiency Treatment Consortium (PIDTC) report. *J Allergy Clin Immunol* 2014; 133(2): 335-347.
20. European Medicines Agency. Strimvelis: European public assessment report; summary for the public [online]. 08.06.2016 [Zugriff: 05.09.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/003854/WC500208202.pdf.
21. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über die Früherkennung von Krankheiten bei Kindern bis zur Vollendung des 6. Lebensjahres (Kinder-Richtlinie) [online]. 07.07.2016 [Zugriff: 11.10.2016]. URL: https://www.g-ba.de/downloads/62-492-1240/RL_Kinder_2015-06-18.pdf.
22. Kubiak C, Jyonouchi S, Kuo C, Garcia-Lloret M, Dorsey MJ, Sleasman J et al. Fiscal implications of newborn screening in the diagnosis of severe combined immunodeficiency. *J Allergy Clin Immunol Pract* 2014; 2(6): 697-702.
23. Kwan A, Church JA, Cowan MJ, Agarwal R, Kapoor N, Kohn DB et al. Newborn screening for severe combined immunodeficiency and T-cell lymphopenia in California: results of the first 2 years. *J Allergy Clin Immunol* 2013; 132(1): 140-150.

24. Brown L, Xu-Bayford J, Allwood Z, Slatter M, Cant A, Davies EG et al. Neonatal diagnosis of severe combined immunodeficiency leads to significantly improved survival outcome: the case for newborn screening. *Blood* 2011; 117(11): 3243-3246.
25. Myers LA, Patel DD, Puck JM, Buckley RH. Hematopoietic stem cell transplantation for severe combined immunodeficiency in the neonatal period leads to superior thymic output and improved survival. *Blood* 2002; 99(3): 872-878.
26. Cuvelier GD, Rubin TS, Wall DA, Schroeder ML. Long-term outcomes of hematopoietic stem cell transplantation for ZAP70 deficiency. *J Clin Immunol* 2016; 36(7): 713-724.
27. Kwan A, Hu D, Song M, Gomes H, Brown DR, Bourque T et al. Successful newborn screening for SCID in the Navajo Nation. *Clin Immunol* 2015; 158(1): 29-34.
28. Vogel BH, Bonagura V, Weinberg GA, Ballow M, Isabelle J, DiAntonio L et al. Newborn screening for SCID in New York State: experience from the first two years. *J Clin Immunol* 2014; 34(3): 289-303.
29. Verbsky JW, Baker MW, Grossman WJ, Hintermeyer M, Dasu T, Bonacci B et al. Newborn screening for severe combined immunodeficiency; the Wisconsin experience (2008-2011). *J Clin Immunol* 2012; 32(1): 82-88.
30. Chien YH, Chiang SC, Chang KL, Yu HH, Lee WI, Tsai LP et al. Incidence of severe combined immunodeficiency through newborn screening in a Chinese population. *J Formos Med Assoc* 2015; 114(1): 12-16.
31. Kanegae MP, Barreiros LA, Mazzucchelli JT, Hadachi SM, De Figueiredo Ferreira Guilhoto LM, Acquesta AL et al. Neonatal screening for severe combined immunodeficiency in Brazil. *J Pediatr (Rio J)* 2016; 92(4): 374-380.
32. De Felipe B, Olbrich P, Lucenas JM, Delgado-Pecellin C, Pavon-Delgado A, Marquez J et al. Prospective neonatal screening for severe T- and B-lymphocyte deficiencies in Seville. *Pediatr Allergy Immunol* 2016; 27(1): 70-77.
33. Olbrich P, De Felipe B, Delgado-Pecellin C, Rodero R, Rojas P, Aguayo J et al. A first pilot study on the neonatal screening of primary immunodeficiencies in Spain: TRECS and KRECS identify severe T- and B-cell lymphopenia [Spanisch]. *An Pediatr (Barc)* 2014; 81(5): 310-317.
34. Nantes University Hospital. Generalized neonatal screening of severe combined immunodeficiencies (DEPISTREC): full text view [online]. In: *ClinicalTrials.gov*. 22.02.2016 [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT02244450>.
35. Primary Immune Deficiency Treatment Consortium. About the PIDTC [online]. [Zugriff: 14.06.2016]. URL: <https://www.rarediseasesnetwork.org/cms/pidtc/About-Us>.
36. United States Immunodeficiency Network. USIDNET Registry: primary immunodeficiency (PI) diseases registry at USIDNET [online]. [Zugriff: 14.06.2016]. URL: <https://usidnet.org/usidnet-registry/>.

37. Laberko A, Gennery AR. Cytoablative conditioning for severe combined immunodeficiency: help or hindrance? *Expert Rev Clin Immunol* 2015; 11(7): 785-788.
38. Cowan MJ, Gennery AR. Radiation-sensitive severe combined immunodeficiency: the arguments for and against conditioning before hematopoietic cell transplantation; what to do? *J Allergy Clin Immunol* 2015; 136(5): 1178-1185.
39. Barbaro M, Ohlsson A, Borte S, Jonsson S, Zetterström R, King J et al. Newborn screening for severe primary immunodeficiency diseases in Sweden: a two-year pilot TREC and KREC screening study [eingereichtes Manuskript]. 2016.
40. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.2. Köln: IQWiG; 2015. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf.
41. 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.
42. 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.
43. 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.
44. 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.
45. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.
46. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003; 138(1): W1-W12.
47. 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.
48. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
49. 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.

50. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
51. 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.
52. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.
53. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the Bernoulli parameter. *Am Stat* 1996; 50(1): 63-68.
54. 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.
55. 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.
56. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-64.
57. Hotelling H. The generalization of student's ratio. *Ann Math Stat* 1931; 2(3): 360-378.
58. Children's Hospital of Fudan University. Neonatal screening of severe combined immunodeficiencies: full text view [online]. In: *ClinicalTrials.gov*. [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT02590328>.
59. Ann & Robert H Lurie Children's Hospital of Chicago. Early diagnosis and stem cell transplantation for severe immunodeficiency diseases (SIDS): full text view [online]. In: *ClinicalTrials.gov*. 26.05.2009 [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT00613561>.
60. National Institute of Allergy and Infectious Diseases. Natural history study of SCID disorders: full text view [online]. In: *ClinicalTrial.gov*. 16.03.2016 [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT01186913>.
61. National Institute of Allergy and Infectious Diseases. Patients treated for SCID (1968-2010): full text view [online]. In: *ClinicalTrials.gov*. 17.03.2016 [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT01346150>.
62. St. Jude Children's Research Hospital. Transplantation of hematopoietic cells in children with severe combined immunodeficiency syndrome: full text view [online]. In: *ClinicalTrials.gov*. 19.05.2009 [Zugriff: 24.03.2016]. URL: <https://ClinicalTrials.gov/show/NCT00152100>.
63. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. *Computat Stat Data Anal* 1994; 17(5): 555-574.

64. Van der Spek J, Groenwold RH, Van der Burg M, Van Montfrans JM. TREC based newborn screening for severe combined immunodeficiency disease: a systematic review. *J Clin Immunol* 2015; 35(4): 416-430.
65. Kane L, Gennery AR, Crooks BN, Flood TJ, Abinun M, Cant AJ. Neonatal bone marrow transplantation for severe combined immunodeficiency. *Arch Dis Child Fetal Neonatal Ed* 2001; 85(2): F110-F113.
66. Antoine C, Muller S, Cant A, Cavazzana-Calvo M, Veys P, Vossen J et al. Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99. *Lancet* 2003; 361(9357): 553-560.
67. Buckley RH, Schiff SE, Schiff RI, Markert L, Williams LW, Roberts JL et al. Hematopoietic stem-cell transplantation for the treatment of severe combined immunodeficiency. *N Engl J Med* 1999; 340(7): 508-516.
68. Buckley RH. Transplantation of hematopoietic stem cells in human severe combined immunodeficiency: longterm outcomes. *Immunol Res* 2011; 49(1-3): 25-43.
69. Chan A, Scalchunes C, Boyle M, Puck JM. Early vs. delayed diagnosis of severe combined immunodeficiency: a family perspective survey. *Clin Immunol* 2011; 138(1): 3-8.
70. Gennery AR, Slatter MA, Grandin L, Taupin P, Cant AJ, Veys P et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? *J Allergy Clin Immunol* 2010; 126(3): 602-10.e1-11.
71. Railey MD, Lokhnygina Y, Buckley RH. Long-term clinical outcome of patients with severe combined immunodeficiency who received related donor bone marrow transplants without pretransplant chemotherapy or post-transplant GVHD prophylaxis. *J Pediatr* 2009; 155(6): 834-840.e1.
72. Comeau AM, Hale JE, Pai SY, Bonilla FA, Notarangelo LD, Pasternack MS et al. Guidelines for implementation of population-based newborn screening for severe combined immunodeficiency. *J Inherit Metab Dis* 2010; 33(Suppl 2): S273-S281.
73. Morinishi Y, Imai K, Nakagawa N, Sato H, Horiuchi K, Ohtsuka Y et al. Identification of severe combined immunodeficiency by T-cell receptor excision circles quantification using neonatal Guthrie cards. *J Pediatr* 2009; 155(6): 829-833.
74. Pai SY, Logan BR, Griffith LM, Buckley RH, Parrott RE, Dvorak CC et al. Transplantation outcomes for severe combined immunodeficiency, 2000-2009. *N Engl J Med* 2014; 371(5): 434-446.
75. Grazioli S, Bennett M, Hildebrand KJ, Vallance H, Turvey SE, Junker AK. Limitation of TREC-based newborn screening for ZAP70 severe combined immunodeficiency. *Clin Immunol* 2014; 153(1): 209-210.

76. Speckmann C, Neumann C, Borte S, Marca GL, Sass JO, Wiech E et al. Delayed-onset adenosine deaminase deficiency: strategies for an early diagnosis. *J Allergy Clin Immunol* 2012; 130(4): 991-994.
77. La Marca G, Canessa C, Giocaliere E, Romano F, Malvagia S, Funghini S et al. Diagnosis of immunodeficiency caused by a purine nucleoside phosphorylase defect by using tandem mass spectrometry on dried blood spots. *J Allergy Clin Immunol* 2014; 134(1): 155-159.
78. Cochrane Library. How CENTRAL is created [online]. [Zugriff: 12.11.2015]. URL: <http://www.cochranelibrary.com/help/central-help.html>.
79. 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: 17.09.2015]. URL: <http://www.dimdi.de/static/de/klassi/icd-10-gm/kodesuche/onlinefassungen/htmlgm2013/index.htm>.

A6 Studienlisten

A6.1 Liste der eingeschlossenen Studien

Vergleichende Interventionsstudien der Screeningkette

Kubiak C, Jyonouchi S, Kuo C, Garcia-Lloret M, Dorsey MJ, Sleasman J et al. Fiscal implications of newborn screening in the diagnosis of severe combined immunodeficiency. *J Allergy Clin Immunol Pract* 2014; 2(6): 697-702.

Vergleichende Interventionsstudien zum Therapiebeginn

Brown L, Xu-Bayford J, Allwood Z, Slatter M, Cant A, Davies EG et al. Neonatal diagnosis of severe combined immunodeficiency leads to significantly improved survival outcome: the case for newborn screening. *Blood* 2011; 117(11): 3243-3246.

Cuvelier GD, Rubin TS, Wall DA, Schroeder ML. Long-term outcomes of hematopoietic stem cell transplantation for ZAP70 deficiency. *J Clin Immunol* 2016; 36(7): 713-724.

Myers LA, Patel DD, Puck JM, Buckley RH. Hematopoietic stem cell transplantation for severe combined immunodeficiency in the neonatal period leads to superior thymic output and improved survival. *Blood* 2002; 99(3): 872-878.

Studien zur diagnostischen Güte

1. Chien YH, Chiang SC, Chang KL, Yu HH, Lee WI, Tsai LP et al. Incidence of severe combined immunodeficiency through newborn screening in a Chinese population. *J Formos Med Assoc* 2015; 114(1): 12-16.

2. De Felipe B, Olbrich P, Lucenas JM, Delgado-Pecellin C, Pavon-Delgado A, Marquez J et al. Prospective neonatal screening for severe T- and B-lymphocyte deficiencies in Seville. *Pediatr Allergy Immunol* 2016; 27(1): 70-77.

3. Kanegae MP, Barreiros LA, Mazzucchelli JT, Hadachi SM, De Figueiredo Ferreira Guilhoto LM, Acquesta AL et al. Neonatal screening for severe combined immunodeficiency in Brazil. *J Pediatr (Rio J)* 2016; 92(4): 374-380.

4. Kwan A, Church JA, Cowan MJ, Agarwal R, Kapoor N, Kohn DB et al. Newborn screening for severe combined immunodeficiency and T-cell lymphopenia in California: results of the first 2 years. *J Allergy Clin Immunol* 2013; 132(1): 140-150.

5. Kwan A, Hu D, Song M, Gomes H, Brown DR, Bourque T et al. Successful newborn screening for SCID in the Navajo Nation. *Clin Immunol* 2015; 158(1): 29-34.

6. Routes JM, Grossman WJ, Verbsky J, Laessig RH, Hoffman GL, Brokopp CD et al. Statewide newborn screening for severe T-cell lymphopenia. *JAMA* 2009; 302(22): 2465-2470.

7. Verbsky JW, Baker MW, Grossman WJ, Hintermeyer M, Dasu T, Bonacci B et al. Newborn screening for severe combined immunodeficiency; the Wisconsin experience (2008-2011). *J Clin Immunol* 2012; 32(1): 82-88.

8. Vogel BH, Bonagura V, Weinberg GA, Ballow M, Isabelle J, DiAntonio L et al. Newborn screening for SCID in New York State: experience from the first two years. *J Clin Immunol* 2014; 34(3): 289-303. Verbsky JW, Baker MW, Grossman WJ, Hintermeyer M, Dasu T, Bonacci B et al. Newborn screening for severe combined immunodeficiency; the Wisconsin experience (2008-2011). *J Clin Immunol* 2012; 32(1): 82-88.

Vogel BH, Bonagura V, Weinberg GA, Ballow M, Isabelle J, DiAntonio L et al. Newborn screening for SCID in New York State: experience from the first two years. *J Clin Immunol* 2014; 34(3): 289-303.

A6.2 Liste der gesichteten systematischen Übersichten

Lipstein EA, Vorono S, Browning MF, Green NS, Kemper AR, Knapp AA et al. Systematic evidence review of newborn screening and treatment of severe combined immunodeficiency. *Pediatrics* 2010; 125(5): e1226-e1235.

Van der Spek J, Groenwold RH, Van der Burg M, Van Montfrans JM. TREC based newborn screening for severe combined immunodeficiency disease: a systematic review. *J Clin Immunol* 2015; 35(4): 416-430.

A6.3 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

Nicht E1 (Population)

1. Antoine C, Muller S, Cant A, Cavazzana-Calvo M, Veys P, Vossen J et al. Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99. *Lancet* 2003; 361(9357): 553-560.

2. Bertrand Y, Landais P, Friedrich W, Gerritsen B, Morgan G, Fasth A et al. Influence of severe combined immunodeficiency phenotype on the outcome of HLA non-identical, T-cell-depleted bone marrow transplantation: a retrospective European survey from the European Group for Bone Marrow Transplantation and the European Society for Immunodeficiency. *J Pediatr* 1999; 134(6): 740-748.

3. Bortin MM, Rimm AA. Severe combined immunodeficiency disease: characterization of the disease and results of transplantation. *JAMA* 1977; 238(7): 591-600.

4. Chan K, Puck JM. Development of population-based newborn screening for severe combined immunodeficiency. *J Allergy Clin Immunol* 2005; 115(2): 391-398.

5. Cipe FE, Dogu F, Aytekin C, Yuksek M, Kendirli T, Yildiran A et al. HLA-haploidentical transplantations for primary immunodeficiencies: a single-center experience. *Pediatr Transplant* 2012; 16(5): 451-457.

6. Fischer A, Griscelli C, Friedrich W, Kubanek B, Levinsky R, Morgan G et al. Bone-marrow transplantation for immunodeficiencies and osteopetrosis: European survey, 1968-1985. *Lancet* 1986; 2(8515): 1080-1084.

7. Fischer A, Landais P, Friedrich W, Morgan G, Gerritsen B, Fasth A et al. European experience of bone-marrow transplantation for severe combined immunodeficiency. *Lancet* 1990; 336(8719): 850-854.
8. Fronkova E, Klocperk A, Svaton M, Novakova M, Kotrova M, Kayserova J et al. The TREC/KREC assay for the diagnosis and monitoring of patients with DiGeorge syndrome. *PLoS One* 2014; 9(12): e114514.
9. Gennery AR, Slatter MA, Grandin L, Taupin P, Cant AJ, Veys P et al. Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: entering a new century, do we do better? *J Allergy Clin Immunol* 2010; 126(3): 602-10.e1-11.
10. Giri N, Vowels M, Ziegler JB, Ford D, Lam-Po-Tang R. HLA non-identical T-cell-depleted bone marrow transplantation for primary immunodeficiency diseases. *Aust N Z J Med* 1994; 24(1): 26-30.
11. Jilkina O, Thompson JR, Kwan L, Van Caesele P, Rockman-Greenberg C, Schroeder ML. Retrospective TREC testing of newborns with severe combined immunodeficiency and other primary immunodeficiency diseases. *Mol Genet Metab Rep* 2014; 1: 324-333.
12. Morio T, Atsuta Y, Tomizawa D, Nagamura-Inoue T, Kato K, Ariga T et al. Outcome of unrelated umbilical cord blood transplantation in 88 patients with primary immunodeficiency in Japan. *Br J Haematol* 2011; 154(3): 363-372.
13. Patel NC, Chinen J, Rosenblatt HM, Hanson IC, Brown BS, Paul ME et al. Long-term outcomes of nonconditioned patients with severe combined immunodeficiency transplanted with HLA-identical or haploidentical bone marrow depleted of T cells with anti-CD6 mAb. *J Allergy Clin Immunol* 2008; 122(6): 1185-1193.
14. Petrovic A, Dorsey M, Miotke J, Shepherd C, Day N. Hematopoietic stem cell transplantation for pediatric patients with primary immunodeficiency diseases at All Children's Hospital/University of South Florida. *Immunol Res* 2009; 44(1-3): 169-178.
15. Railey MD, Lokhnygina Y, Buckley RH. Long-term clinical outcome of patients with severe combined immunodeficiency who received related donor bone marrow transplants without pretransplant chemotherapy or post-transplant GVHD prophylaxis. *J Pediatr* 2009; 155(6): 834-840.e1.
16. Smogorzewska EM, Brooks J, Annett G, Kapoor N, Crooks GM, Kohn DB et al. T cell depleted haploidentical bone marrow transplantation for the treatment of children with severe combined immunodeficiency. *Arch Immunol Ther Exp (Warsz)* 2000; 48(2): 111-118.
17. Teigland CL, Parrott RE, Buckley RH. Long-term outcome of non-ablative booster BMT in patients with SCID. *Bone Marrow Transplant* 2013; 48(8): 1050-1055.
18. Titman P, Pink E, Skucek E, O'Hanlon K, Cole TJ, Gaspar J et al. Cognitive and behavioral abnormalities in children after hematopoietic stem cell transplantation for severe congenital immunodeficiencies. *Blood* 2008; 112(9): 3907-3913.

19. Tsuji Y, Imai K, Kajiwara M, Aoki Y, Isoda T, Tomizawa D et al. Hematopoietic stem cell transplantation for 30 patients with primary immunodeficiency diseases: 20 years experience of a single team. *Bone Marrow Transplant* 2006; 37(5): 469-477.

Nicht E2 (Intervention / Indextest)

1. Al-Ghoniaim A. Stem cell transplantation for primary immunodeficiencies: King Faisal Specialist Hospital experience from 1993 to 2006. *Bone Marrow Transplant* 2008; 42(Suppl 1): S53-S56.

2. Al-Herz W, Notarangelo LD, Sadek A, Buckley R. Combined immunodeficiency in the United States and Kuwait: Comparison of patients' characteristics and molecular diagnosis. *Clin Immunol* 2015; 161(2): 170-173.

3. Azzari C, La Marca G, Resti M. Neonatal screening for severe combined immunodeficiency caused by an adenosine deaminase defect: a reliable and inexpensive method using tandem mass spectrometry. *J Allergy Clin Immunol* 2011; 127(6): 1394-1399.

4. Bousfiha AA, Jeddane L, El Hafidi N, Benajiba N, Rada N, El Bakkouri J et al. First report on the Moroccan registry of primary immunodeficiencies: 15 years of experience (1998-2012). *J Clin Immunol* 2014; 34(4): 459-468.

5. Chan A, Scalchunes C, Boyle M, Puck JM. Early vs. delayed diagnosis of severe combined immunodeficiency: a family perspective survey. *Clin Immunol* 2011; 138(1): 3-8.

6. De Pagter AP, Bredius RG, Kuijpers TW, Tramper J, Van der Burg M, Van Montfrans J et al. Overview of 15-year severe combined immunodeficiency in the Netherlands: towards newborn blood spot screening. *Eur J Pediatr* 2015; 174(9): 1183-1188.

7. Dinardo L, Brown V, Perez E, Bunin N, Sullivan KE. A single-center study of hematopoietic stem cell transplantation for primary immune deficiencies (PIDD). *Pediatr Transplant* 2012; 16(1): 63-72.

8. Dvorak CC, Cowan MJ, Logan BR, Notarangelo LD, Griffith LM, Puck JM et al. The natural history of children with severe combined immunodeficiency: baseline features of the first fifty patients of the primary immune deficiency treatment consortium prospective study 6901. *J Clin Immunol* 2013; 33(7): 1156-1164.

9. Eapen M, Ahn KW, Orchard PJ, Cowan MJ, Davies SM, Fasth A et al. Long-term survival and late deaths after hematopoietic cell transplantation for primary immunodeficiency diseases and inborn errors of metabolism. *Biol Blood Marrow Transplant* 2012; 18(9): 1438-1445.

10. Haddad E, Landais P, Friedrich W, Gerritsen B, Cavazzana-Calvo M, Morgan G et al. Long-term immune reconstitution and outcome after HLA-nonidentical T-cell-depleted bone marrow transplantation for severe combined immunodeficiency: a European retrospective study of 116 patients. *Blood* 1998; 91(10): 3646-3653.

11. Hamidieh AA, Behfar M, Pourpak Z, Faghihi-Kashani S, Fazlollahi MR, Hosseini AS et al. Long-term outcomes of fludarabine, melphalan and antithymocyte globulin as reduced-intensity conditioning regimen for allogeneic hematopoietic stem cell transplantation in children with primary immunodeficiency disorders: a prospective single center study. *Bone Marrow Transplant* 2016; 51(2): 219-226.
12. La Marca G, Canessa C, Giocaliere E, Romano F, Duse M, Malvagia S et al. Tandem mass spectrometry, but not T-cell receptor excision circle analysis, identifies newborns with late-onset adenosine deaminase deficiency. *J Allergy Clin Immunol* 2013; 131(6): 1604-1610.
13. La Marca G, Giocaliere E, Malvagia S, Funghini S, Ombrone D, Della Bona ML et al. The inclusion of ADA-SCID in expanded newborn screening by tandem mass spectrometry. *J Pharm Biomed Anal* 2014; 88: 201-206.
14. Lim DL, Thong BY, Ho SY, Shek LP, Lou J, Leong KP et al. Primary immunodeficiency diseases in Singapore: the last 11 years. *Singapore Med J* 2003; 44(11): 579-586.
15. Mazzolari E, De Martiis D, Forino C, Lanfranchi A, Giliani S, Marzollo R et al. Single-center analysis of long-term outcome after hematopoietic cell transplantation in children with congenital severe T cell immunodeficiency. *Immunol Res* 2009; 44(1-3): 4-17.
16. Mellouli F, Mustapha IB, Khaled MB, Besbes H, Ouederni M, Mekki N et al. Report of the tunisian registry of primary immunodeficiencies: 25-years of experience (1988-2012). *J Clin Immunol* 2015; 35(8): 745-753.
17. Mendez-Echevarria A, Del Rosal T, Perez-Costa E, Rodriguez-Pena R, Zarauza A, Ferreira-Cerdan A et al. Clinical features before hematopoietic stem cell transplantation or enzyme replacement therapy of children with combined immunodeficiency. *Pediatr Infect Dis J* 2016; 35(7): 794-798.
18. Meshaal S, El Hawary R, Elsharkawy M, Mousa RK, Farid RJ, Abd Elaziz D et al. Mutations in Recombination Activating Gene 1 and 2 in patients with severe combined immunodeficiency disorders in Egypt. *Clin Immunol* 2015; 158(2): 167-173.
19. Mitchell R, Nivison-Smith I, Anazodo A, Tiedemann K, Shaw P, Teague L et al. Outcomes of hematopoietic stem cell transplantation in primary immunodeficiency: a report from the Australian and New Zealand Children's Haematology Oncology Group and the Australasian Bone Marrow Transplant Recipient Registry. *Biol Blood Marrow Transplant* 2013; 19(3): 338-343.
20. Neven B, Leroy S, Decaluwe H, Le Deist F, Picard C, Moshous D et al. Long-term outcome after hematopoietic stem cell transplantation of a single-center cohort of 90 patients with severe combined immunodeficiency. *Blood* 2009; 113(17): 4114-4124.
21. Odek C, Kendirli T, Dogu F, Yaman A, Vatansever G, Cipe F et al. Patients with primary immunodeficiencies in pediatric intensive care unit: outcomes and mortality-related risk factors. *J Clin Immunol* 2014; 34(3): 309-315.

22. Olbrich P, De Felipe B, Delgado-Pecellin C, Rodero R, Rojas P, Aguayo J et al. A first pilot study on the neonatal screening of primary immunodeficiencies in Spain: TRECS and KRECS identify severe T- and B-cell lymphopenia [Spanisch]. *An Pediatr (Barc)* 2014; 81(5): 310-317.
23. Pai SY, Logan BR, Griffith LM, Buckley RH, Parrott RE, Dvorak CC et al. Transplantation outcomes for severe combined immunodeficiency, 2000-2009. *N Engl J Med* 2014; 371(5): 434-446.
24. Pasic S, Vujic D, Veljkovic D, Slavkovic B, Mostarica-Stojkovic M, Minic P et al. Severe combined immunodeficiency in Serbia and Montenegro between years 1986 and 2010: a single-center experience. *J Clin Immunol* 2014; 34(3): 304-308.
25. Rao K, Adams S, Qasim W, Allwood Z, Worth A, Silva J et al. Effect of stem cell source on long-term chimerism and event-free survival in children with primary immunodeficiency disorders after fludarabine and melphalan conditioning regimen. *J Allergy Clin Immunol* 20.04.2016 [Epub ahead of print].
26. Rousso SZ, Shamriz O, Zilkha A, Braun J, Averbuch D, Or R et al. Hematopoietic stem cell transplantations for primary immune deficiencies: 3 decades of experience from a tertiary medical center. *J Pediatr Hematol Oncol* 2015; 37(5): e295-e300.
27. Rozmus J, Junker A, Thibodeau ML, Grenier D, Turvey SE, Yacoub W et al. Severe combined immunodeficiency (SCID) in Canadian children: a national surveillance study. *J Clin Immunol* 2013; 33(8): 1310-1316.
28. Stephan JL, Vlekova V, Le Deist F, Blanche S, Donadieu J, De Saint-Basile G et al. Severe combined immunodeficiency: a retrospective single-center study of clinical presentation and outcome in 117 patients. *J Pediatr* 1993; 123(4): 564-572.
29. Wahlstrom J, Patel K, Eckhert E, Kong D, Horn B, Cowan MJ et al. Transplacental maternal engraftment and posttransplantation graft-versus-host disease in children with severe combined immunodeficiency. *J Allergy Clin Immunol* 16.06.2016 [Epub ahead of print].
30. Yao CM, Han XH, Zhang YD, Zhang H, Jin YY, Cao RM et al. Clinical characteristics and genetic profiles of 44 patients with severe combined immunodeficiency (SCID): report from Shanghai, China (2004-2011). *J Clin Immunol* 2013; 33(3): 526-539.

Nicht E3 (Vergleich / Referenztest)

1. Adams SP, Rashid S, Premachandra T, Harvey K, Ifederu A, Wilson MC et al. Screening of neonatal UK dried blood spots using a duplex TREC screening assay. *J Clin Immunol* 2014; 34(3): 323-330.
2. Shearer WT, Dunn E, Notarangelo LD, Dvorak CC, Puck JM, Logan BR et al. Establishing diagnostic criteria for severe combined immunodeficiency disease (SCID), leaky SCID, and Omenn syndrome: the Primary Immune Deficiency Treatment Consortium experience. *J Allergy Clin Immunol* 2014; 133(4): 1092-1098.

Nicht E4 (patientenrelevante Endpunkte / Zielgrößen)

1. Naylor EW, Orfanos AP, Guthrie R. An improved screening test for adenosine deaminase deficiency. *J Pediatr* 1978; 93(3): 473-476.

Nicht E5 (Studientyp)

1. Akar HH, Patiroglu T, Hershfield M, Van der Burg M. Combined immunodeficiencies: twenty years experience from a single center in Turkey. *Cent Eur J Immunol* 2016; 41(1): 107-115.

2. Baker MW, Grossman WJ, Laessig RH, Hoffman GL, Brokopp CD, Kurtycz DF et al. Development of a routine newborn screening protocol for severe combined immunodeficiency. *J Allergy Clin Immunol* 2009; 124(3): 522-527.

3. Borte S, Von Döbeln U, Fasth A, Wang N, Janzi M, Winiarski J et al. Neonatal screening for severe primary immunodeficiency diseases using high-throughput triplex real-time PCR. *Blood* 2012; 119(11): 2552-2555.

4. Comeau AM, Hale JE, Pai SY, Bonilla FA, Notarangelo LD, Pasternack MS et al. Guidelines for implementation of population-based newborn screening for severe combined immunodeficiency. *J Inher Metab Dis* 2010; 33(Suppl 2): S273-S281.

5. Ding Y, Thompson JD, Kobrynski L, Ojodu J, Zarbalian G, Grosse SD. Cost-effectiveness/cost-benefit analysis of newborn screening for severe combined immune deficiency in Washington State. *J Pediatr* 2016; 172: 127-135.

6. Edgar JDM, Buckland M, Guzman D, Conlon NP, Knerr V, Bangs C et al. The United Kingdom Primary Immune Deficiency (UKPID) Registry: report of the first 4 years' activity 2008-2012. *Clin Exp Immunol* 2014; 175(1): 68-78.

7. Kwan A, Abraham RS, Currier R, Brower A, Andruszewski K, Abbott JK et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. *JAMA* 2014; 312(7): 729-738.

8. Liebrich W, Emanuel S, Esser M. Primary immunodeficiency. *Current allergy & clinical immunology* 2015; 28(4): 305-307.

9. McGhee SA, Stiehm ER, Cowan M, Krogstad P, McCabe ER. Two-tiered universal newborn screening strategy for severe combined immunodeficiency. *Mol Genet Metab* 2005; 86(4): 427-430.

10. McWilliams LM, Dell Railey M, Buckley RH. Positive family history, infection, low absolute lymphocyte count (ALC), and absent thymic shadow: diagnostic clues for all molecular forms of severe combined immunodeficiency (SCID). *J Allergy Clin Immunol Pract* 2015; 3(4): 585-591.

11. Morinishi Y, Imai K, Nakagawa N, Sato H, Horiuchi K, Ohtsuka Y et al. Identification of severe combined immunodeficiency by T-cell receptor excision circles quantification using neonatal Guthrie cards. *J Pediatr* 2009; 155(6): 829-833.

12. Muller SM, Friedrich W. Stem cell transplantation for treatment of primary immunodeficiency disorders. *Iran J Allergy Asthma Immunol* 2005; 4(1): 1-8.
13. Nicholas S, Krance RA, Hanson IC, Chinen J, Mamlök RJ, Roifman CM et al. Early versus delayed diagnosis of SCID: triumph versus tragedy. *Clin Immunol* 2011; 139(3): 360-362.
14. Papiroglu T, Akar HH, Unal E, Ozdemir MA, Karakukcu M. Hematopoietic stem cell transplant for primary immunodeficiency diseases: a single-center experience. *Exp Clin Transplant* 21.03.2016 [Epub ahead of print].
15. Somech R, Lev A, Simon AJ, Korn D, Garty BZ, Amariglio N et al. Newborn screening for severe T and B cell immunodeficiency in Israel: a pilot study. *Isr Med Assoc J* 2013; 15(8): 404-409.

Nicht E6 (Vollpublikation)

1. Azzari C. Neonatal screening of severe combined immunodeficiency [Italienisch]. *Area Pediatrica* 2012; 13(1): 20-24.
2. Buckley RH. Transplantation of hematopoietic stem cells in human severe combined immunodeficiency: longterm outcomes. *Immunol Res* 2011; 49(1-3): 25-43.
3. Kwan A, Abraham RS, Currier R, Brower A, Andruszewski K, Abbott JK et al. Erratum: "Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States" (*JAMA* 2014; 312(7): 729-738). *JAMA* 2014; 312(12): 1259.
4. Leinmüller R. Newborn screening for severe immunodeficiencies: early diagnosis correlates with high cure rates. *Dtsch Arztebl Int* 2015; 112(1-2): A31-A32.
5. Thomas C, Mirallié S, Pierres C, Dert C, Clément MC, Mahlaoui N et al. Neonatal screening of severe combined immunodeficiencies [Französisch]. *Arch Pediatr* 2015; 22(6): 646-652.
6. Verbsky J, Thakar M, Routes J. The Wisconsin approach to newborn screening for severe combined immunodeficiency. *J Allergy Clin Immunol* 2012; 129(3): 622-627.

A7 Suchstrategien

A7.1 Suchstrategien in bibliografischen Datenbanken

A7.1.1 Embase

Suchoberfläche: Ovid

- Embase 1974 to 2016 August 12

Es wurden folgende Filter übernommen:

- Ausschluss von Tierstudien: Cochrane Library [78]

#	Searches
1	*combined immunodeficiency/
2	exp severe combined immunodeficiency/
3	((primary or severe) adj3 (immunodeficienc* or T-cell lymphopenia)).ab,ti.
4	(scid not (mouse or mice)).mp.
5	or/1-4
6	exp infant/
7	(newborn* or neonat* or infan* or pediatric*).ti,ab.
8	or/6-7
9	Newborn screening/
10	and/5,9
11	screening*.ab,ti.
12	T lymphocyte receptor/
13	((kappa or T-cell) adj5 excision circle*).ab,ti.
14	Interleukin 7/
15	(interleukin-7 or IL-7).ab,ti.
16	or/11-15
17	(and/5,8,16) or 10
18	exp stem cell transplantation/
19	exp bone marrow transplantation/
20	((bone marrow* or stem cell* or cord blood or hematopoietic) adj3 transplant*).ab,ti.
21	or/18-20
22	and/5,21
23	or/17,22
24	23 not medline*.cr.
25	case report/ not exp controlled study/
26	24 not 25

#	Searches
27	exp experimental organism/ or animal tissue/ or animal cell/ or exp animal disease/ or exp carnivore disease/ or exp bird/ or exp experimental animal welfare/ or exp animal husbandry/ or animal behavior/ or exp animal cell culture/ or exp mammalian disease/ or exp mammal/ or exp marine species/ or nonhuman/ or animal.hw.
28	27 not human/
29	26 not 28

A7.1.2 MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations August 12, 2016
- Ovid MEDLINE(R) 1946 to August Week 1 2016
- Ovid MEDLINE(R) Daily Update August 12, 2016
- Ovid MEDLINE(R) Epub Ahead of Print August 12, 2016

#	Searches
1	Severe Combined Immunodeficiency/
2	Immunologic Deficiency Syndromes/su
3	((primary or severe) adj3 (immunodeficienc* or T-cell lymphopenia)).ab,ti.
4	(scid not (mouse or mice)).mp.
5	or/1-4
6	(newborn* or neonat* or infan* or pediatric*).ti,ab.
7	exp infant/
8	or/6-7
9	Neonatal Screening/
10	and/5,9
11	screening*.ab,ti.
12	Receptors, Antigen, T-Cell/
13	((kappa or T-cell) adj5 excision circle*).ab,ti.
14	Interleukin-7/
15	(interleukin-7 or IL-7).ab,ti.
16	or/11-15
17	(and/5,8,16) or 10
18	exp Stem Cell Transplantation/

#	Searches
19	Bone Marrow Transplantation/
20	((bone marrow* or stem cell* or cord blood or hematopoietic) adj3 transplant*).ab,ti.
21	or/18-20
22	and/5,21
23	or/17,22
24	animals/ not humans/
25	23 not 24

A7.1.3 PubMed

Suchoberfläche: NLM

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

Search	Query
#1	Search ((primary [TIAB] OR severe [TIAB]) AND (immunodeficienc* [TIAB] OR T-cell lymphopenia [TIAB]))
#2	Search (SCID NOT (mouse or mice))
#3	Search (#1 OR #2)
#4	Search (newborn*[TIAB] OR neonat*[TIAB] OR infan*[TIAB] OR pediatric*[TIAB])
#5	Search screening* [TIAB]
#6	Search ((kappa [TIAB] OR T-cell [TIAB]) AND excision circle* [TIAB])
#7	Search (interleukin-7 [TIAB] OR IL-7 [TIAB])
#8	Search (#5 OR #6 OR #7)
#9	Search (#3 AND #4 AND #8)
#10	Search ((bone marrow* [TIAB] OR stem cell* [TIAB] OR cord blood [TIAB] OR hematopoietic [TIAB]) AND transplant* [TIAB])
#11	Search (#3 AND #10)
#12	Search (#9 OR #11)
#14	Search (#12 NOT medline[sb])

A7.1.4 The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews: Issue 8 of 12, August 2016
- Database of Abstracts of Reviews of Effect: Issue 2 of 4, April 2015
- Cochrane Central Register of Controlled Trials: Issue 7 of 12, July 2016
- Health Technology Assessment Database: Issue 3 of 4, July 2016

ID	Search
#1	[mh "Severe Combined Immunodeficiency "]
#2	[mh "Immunologic Deficiency Syndromes "/SU]
#3	((primary or severe) near/3 (immunodeficienc* or T-cell lymphopenia))
#4	("SCID" not (mouse or mice)) .ab,ti
#5	{or #1-#4}
#6	newborn* or neonat* or infan* or pediatric*
#7	[mh infant]
#8	[79-#7]
#9	[mh "Neonatal Screening "]
#10	{and #5, #9}
#11	screening*
#12	[mh "Receptors, Antigen, T-Cell "]
#13	((kappa or T-cell) near/5 excision circle*)
#14	[mh Interleukin-7]
#15	(interleukin-7 or IL-7)
#16	#11 OR #12 OR #13 OR #14 OR #15
#17	{and #5, #8, #16} or #10
#18	[mh "Stem Cell Transplantation "]
#19	[mh "Bone Marrow Transplantation "]
#20	((bone marrow* or stem cell* or cord blood or hematopoietic) near/3 transplant*)
#21	{or #18-#20}
#22	{and #5, #21}
#23	{or #17, #22} in Cochrane Reviews (Reviews and Protocols), Other Reviews, Trials and Technology Assessments

A7.2 Suche in Studienregistern

A7.2.1 ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

URL: <http://www.clinicaltrials.gov>

Eingabeoberfläche: Advanced Search

Suchstrategie

((primary OR severe) AND (immunodeficiency OR T-cell lymphopenia) OR SCID)
AND ((screening OR kappa excision circle OR T-cell excision circle OR interleukin-7 OR
IL-7) AND (Infant OR newborn OR neonatal OR pediatric) OR transplantation)

A7.2.2 EU Clinical Trials Register

Anbieter: European Medicines Agency

URL: <https://www.clinicaltrialsregister.eu>

Eingabeoberfläche: Basic Search

Suchstrategie

("primary immunodeficiency" OR "severe immunodeficiency" OR "primary T-cell
lymphopenia" OR "severe T-cell lymphopenia" OR "severe combined immunodeficiency")
AND (screening OR "kappa excision circle" OR "T-cell excision circle" OR interleukin-7
OR IL-7 OR transplant*)

A7.2.3 International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

URL: <http://apps.who.int/trialsearch/>

Eingabeoberfläche: Advanced Search

Suchstrategie

primary immunodeficiency OR severe immunodeficiency OR primary T-cell lymphopenia
OR severe T-cell lymphopenia OR severe combined immunodeficiency OR SCID *in the
Condition*
Recruitment status is ALL

A8 Offenlegung potenzieller Interessenkonflikte

A8.1 Darlegung potenzieller Interessenkonflikte der externen Sachverständigen

Im Folgenden sind die potenziellen Interessenkonflikte der externen Sachverständigen dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“ mit Stand 11/2013. Das aktuelle Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Becker, Monika	nein	nein	nein	ja	nein	nein
Borte, Michael	nein	ja	ja	ja	ja	nein
Mathes, Tim	nein	nein	nein	ja	nein	nein
Pieper, Dawid	nein	nein	nein	ja	nein	nein

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version 11/2013):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung², für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung², für die Sie tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

² Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.