

Früherkennung eines Vitamin-B12-Mangels und weiterer Zielerkrankungen (Homocystinurie, Propionazidämie und Methylmalonazidurie) im erweiterten Neugeborenenenscreening (ENS)

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The word 'BERICHTSPLAN' is written in white capital letters on a dark blue segment that spans across the middle of the bar.

BERICHTSPLAN

Projekt: S22-02

Version: 1.0

Stand: 13.03.2023

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Früherkennung eines Vitamin-B12-Mangels und weiterer Zielerkrankungen (Homocystinurie, Propionazidämie und Methylmalonazidurie) im erweiterten Neugeborenenenscreening (ENS)

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

24.11.2022

Interne Projektnummer

S22-02

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

Schlagwörter

Neugeborenencreening, Vitamin-B12-Mangel, Homocystinurie, Propionazidämie, Methylmalonazidurie, Nutzenbewertung, Systematische Übersicht

Keywords

Neonatal Screening, Vitamin B12 Deficiency, Homocystinuria, Propionic Acidemia, Methylmalonic Acidemia, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis.....	vi
1 Hintergrund.....	1
2 Fragestellung.....	4
3 Projektverlauf.....	5
4 Methoden	6
4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette	6
4.1.1 Population	6
4.1.2 Prüf- und Vergleichsintervention	6
4.1.3 Patientenrelevante Endpunkte	7
4.1.4 Studientypen	7
4.1.5 Studiendauer	8
4.1.6 Publikationssprache	8
4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	8
4.2 Kriterien für den Einschluss von vergleichenden Interventionsstudien zu Therapiebeginn	9
4.2.1 Population	9
4.2.2 Prüf- und Vergleichsintervention	10
4.2.3 Patientenrelevante Endpunkte	10
4.2.4 Studientypen	10
4.2.5 Studiendauer	11
4.2.6 Publikationssprache	11
4.2.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	11
4.3 Kriterien für den Einschluss von Studien zur diagnostischen Güte.....	12
4.3.1 Population	12
4.3.2 Indextest.....	12
4.3.3 Referenztest	13
4.3.4 Zielgrößen.....	13
4.3.5 Studientypen	13
4.3.6 Studiendauer	13
4.3.7 Publikationssprache	13

4.3.8	Tabellarische Darstellung der Kriterien für den Studieneinschluss	14
4.4	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	14
4.5	Informationsbeschaffung.....	15
4.5.1	Fokussierte Informationsbeschaffung von systematischen Übersichten	15
4.5.2	Umfassende Informationsbeschaffung von Studien	15
4.5.3	Anwendung von Limitierungen auf Datenbankebene	16
4.5.4	Selektion relevanter Studien aus der umfassenden Informationsbeschaffung... ..	17
4.6	Informationsbewertung und -synthese	17
4.6.1	Darstellung der Einzelstudien.....	17
4.6.1.1	Darstellung der Einzelstudien (vergleichende Interventionsstudien)	17
4.6.1.2	Darstellung der Einzelstudien (Studien zur diagnostischen Güte)	18
4.6.2	Bewertung des Verzerrungspotenzials der Ergebnisse.....	18
4.6.2.1	Bewertung des Verzerrungspotenzials der Ergebnisse (vergleichende Interventionsstudien)	18
4.6.2.2	Bewertung des Verzerrungspotenzials der Ergebnisse (Studien zur diagnostischen Güte)	19
4.6.3	Metaanalysen	19
4.6.3.1	Metaanalysen (vergleichende Interventionsstudien)	19
4.6.3.2	Metaanalysen (Studien zur diagnostischen Güte).....	21
4.6.4	Sensitivitätsanalysen	21
4.6.5	Subgruppenmerkmale und andere Effektmodifikatoren.....	22
4.6.6	Aussagen zur Beleglage	22
5	Literatur	24
Anhang A	Suchstrategien in bibliografischen Datenbanken.....	28
A.1	Bibliografische Datenbanken	28

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)	9
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)	12
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)	14
Tabelle 4: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	23

Abkürzungsverzeichnis

Abkürzung	Bedeutung
cbl	Cobalamin
CBS	Cystathionin- β -Synthase
CoA	Coenzym-A
ENS	erweitertes Neugeborenencreening
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
HU	Homocystinurie
MMA	Methylmalonazidurie
MTHFR	Methylentetrahydrofolat-Reduktase
mut0	vollständiger Ausfall der Methylmalonyl-CoA-Mutase
mut-	partieller Ausfall der Methylmalonyl-CoA-Mutase
PA	Propionazidämie
PPV	positiver prädiktiver Wert
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
SÜ	systematische Übersicht
VOPT	Verification of only positive Testers

1 Hintergrund

Vitamin-B12-Mangel, Homocystinurie, Propionazidämie und auch Methylmalonazidurie können die motorische und neurokognitive Entwicklung von Kindern gefährden. In Abhängigkeit von der zugrunde liegenden Erkrankung kann das Leben eines Neugeborenen in Gefahr sein.

Ein Vitamin-B12-Mangel bei Neugeborenen kann durch einen mütterlichen Vitamin-B12-Mangel [1] oder durch die seltene angeborene (autosomal-rezessiv vererbte) genetische Erkrankung Transcobalamin-II-Mangel [2] bedingt sein. Vitamin-B12 ist ein wichtiger Co-Faktor für Stoffwechselprozesse, insbesondere im Zentralnervensystem; ein Mangel kann zu irreversiblen neurologischen Schäden und Entwicklungsstörungen führen [1,2]. Der sensitivste Parameter zum Finden eines Vitamin-B12-Mangels ist die Methylmalonsäure im Urin, welche erhöht ist; der Homocysteinwert im Plasma ist meist auch erhöht [1,2]. Eine genetische Analyse ermöglicht den Nachweis eines Transcobalamin-II-Mangels [2]. Die Angaben zur Häufigkeit des mütterlich bedingten Vitamin-B12-Mangels variieren zwischen 1:5300 [1] bis zu weniger als 1:100 000 Neugeborene [3] und resultieren im Wesentlichen aus Untersuchungen in Neugeborenen Screenings auf eben einen solchen Mangel. Für den Transcobalamin-II-Mangel wird auf der Seite <https://www.orpha.net>, dem Portal für seltene Erkrankungen und Orphan Drugs, eine Prävalenz von weniger als 1:1 000 000 angegeben [4]. Bei beiden Mangelzuständen sollte durch parenterale Gabe (intravenös, intramuskulär) von Vitamin-B12 schnellstmöglich behandelt werden [1]. Während der mütterlich bedingte Mangel nur kurzzeitig behandelt werden muss [1], ist bei der genetisch bedingten Erkrankung eine lebenslange parenterale Supplementierung erforderlich [2].

Bei der Homocystinurie (HU) handelt es sich um eine Sammelbezeichnung für mehrere angeborene (autosomal-rezessiv vererbte) Stoffwechselstörungen, denen eine Anreicherung der Aminosäure Homocystein im Blut bzw. Homocystin im Urin gemein ist [5]. Erhöhte Homocysteinwerte sowie andere sich anstauende Intermediate des Homocysteinstoffwechsels vermitteln die Symptome [6]. Der klassischen HU, auch als Typ-1-HU bezeichnet, liegt ein Cystathionin- β -Synthase(CBS)-Mangel zugrunde [7]. Die Erkrankung manifestiert sich an Augen, Skelett, Gefäßsystem und nicht zuletzt im Gehirn [7,8]. Der Nachweis der Erkrankung erfolgt über eine genetische Analyse [9]. Die Angaben zur ihrer Häufigkeit bei Neugeborenen variieren zwischen 1:1800 und 1:900 000 [10]. Die lebenslange Therapie beinhaltet im Wesentlichen die Gabe von Vitaminen sowie ggf. eine eiweißarme Diät [8]. Auch die angeborenen Cobalamin(cbl)-Stoffwechseldefekte cblD1, cblE sowie cblG, die jeweils Subtypen eines Methylcobalaminmangels bedingen, werden unter HU, auch bezeichnet als HU ohne Methylmalonazidurie, zusammengefasst [11]. Sie manifestieren sich insbesondere im Gehirn; im Rahmen banaler Infekte kann es zum metabolischen Schlaganfall („metabolic stroke“) mit irreversiblen neurologischen Schäden kommen [6,11]. Der Nachweis des jeweils zugrunde liegenden Defekts erfolgt über eine genetische Analyse [11]. Angaben zur Häufigkeit

dieser Stoffwechselstörungen liegen nicht vor. Die lebenslange Therapie beinhaltet im Wesentlichen die Gabe von Vitamin-B12 und ggf. von weiteren Vitaminen [6]. Als dritte Form der HU wird die Hyperhomocysteinämie, bedingt durch einen autosomal-rezessiv vererbten Methylentetrahydrofolat-Reduktase(MTHFR)-Mangel (auch als Typ-2-HU bezeichnet) beschrieben. Oft manifestiert sich der MTHFR-Mangel mit schweren neurologischen Symptomen [6] bis hin zum Koma und Tod [12]. Der Nachweis der Erkrankung erfolgt über eine genetische Analyse [6]. Angaben zur Häufigkeit des MTHFR-Mangels liegen nicht vor. Wesentlicher Baustein der Therapie ist die lebenslange Gabe von hochdosiertem Betain [6,12].

Die Propionazidämie (PA) ist eine seltene angeborene (autosomal-rezessiv vererbte) Stoffwechselstörung und gehört zur Gruppe der Organoazidopathien [5]. Bedingt durch die eingeschränkte Aktivität der Propionyl-Coenzym-A(CoA)-Carboxylase kommt es zu einer Anreicherung von Propionsäure, die regelhaft in unterschiedlichen Stoffwechselprozessen anfällt und bei Katabolie vermehrt gebildet wird [13,14]. Dies kann zu schweren Stoffwechselentgleisungen führen, die unbehandelt zu irreversiblen Gehirnschäden bis hin zum Tod führen (metabolischer Schlaganfall) [5,13,14]. Die Diagnose setzt einen erhöhten Spiegel an organischen Säuren im Urin und Propionylcarnitin im Blut voraus; die Konfirmationsdiagnostik erfolgt über eine genetische Analyse [13,15]. Die Angaben zur Häufigkeit der PA variieren zwischen mehr als 1:50 000 [5] und weniger als 1:100 000 [15]. Während die Akuttherapie in der Behandlung der Stoffwechselentgleisung und damit der Sicherung des Überlebens besteht [13], basiert die Langzeittherapie auf einer strengen bewussten Diät, der regelmäßigen Einnahme von Carnitin zwecks Stabilisierung der Stoffwechsellage [5,13,14] sowie Maßnahmen zur Minimierung der Propionsäurebildung durch Bakterien im Darm [5,14]. Zur Stabilisierung der Stoffwechsellage wird zudem eine Lebertransplantation im Kleinkindalter diskutiert [13,14]. Darüber hinaus ist eine Kontrolle und Therapie möglicher (Langzeit-)Schäden, beispielsweise hinsichtlich neurologischer, nephrologischer, kardiologischer sowie pankreatischer Komplikationen indiziert [13,14].

Bei der Methylmalonazidurie (MMA) handelt es sich um eine Sammelbezeichnung für mehrere seltene angeborene (autosomal-rezessiv vererbte) Stoffwechselstörungen, die wie die PA zur Gruppe der Organoazidopathien gehören [5]. Die Häufigkeit einer MMA wird mit 1:250 000 angegeben [5]. Zu unterscheiden ist die MMA ohne HU [16] von einer MMA mit HU [17]. Dabei ist die MMA ohne HU nochmals nach Vitamin-B12-responsiv [18] bzw. Vitamin-B12-resistent Typ mut0 [19] und Vitamin-B12-resistent Typ mut- [20] zu unterscheiden.

Der Vitamin-B12-sensiblen MMA ohne HU liegen unterschiedliche cbl-Defekte (cblA, cblB oder cblD2) zugrunde, die jeweils einen Adenosylcobalaminmangel bedingen [18]. Stoffwechselentgleisungen bis hin zu Koma sowie Schädigungen des Gehirns und deren Folgen sind zentrale Symptome der Erkrankung; Langzeitfolgen betreffen neben dem Gehirn vor

allein die Nieren [18]. Der Nachweis der Erkrankung erfolgt über eine genetische Analyse [13]. Während die Akuttherapie in der Behandlung der Stoffwechsellage und damit der Sicherung des Überlebens besteht, basiert die Langzeittherapie auf einer proteinarmen Diät (supplementiert mit Spezialaminosäuremischungen), der Gabe von Vitamin-B12 und ggf. von Carnitin zwecks Stabilisierung der Stoffwechsellage [13,18] sowie Maßnahmen zur Minimierung der Propionsäurebildung durch Bakterien im Darm [18]. Abgesehen von der Vitamin-B12-Gabe stellt sich die Langzeittherapie von Vitamin-B12-resistenter MMA ohne HU vergleichbar dar [21], im Fall der Vitamin-B12-resistenten MMA ohne HU sind jedoch auch Fälle bekannt, die auf die Gabe von Vitamin-B12 ansprechen [20]. Aufgrund der Schwere der Erkrankung kann zudem eine Lebertransplantation in Erwägung gezogen werden [13,14,21]. Dabei weisen Betroffene mit einem vollständigen Ausfall der Methylmalonyl-CoA-Mutase (mut0) gegenüber denjenigen mit einem partiellen Enzymausfall (mut-) die schwerwiegenderen Verläufe sowohl in der Akutphase als auch im Langzeitverlauf auf [21]. Die Diagnose der Erkrankung beinhaltet eine Bestimmung organischer Säuren im Urin und Propionylcarnitin im Blut, gefolgt von einer genetischen Analyse [13].

Der MMA mit HU liegen unterschiedliche cbl-Defekte (cblC, cblD, cblF oder cblJ) zugrunde, die jeweils gleichzeitig sowohl einen Methylcobalamin- als auch Adenosylcobalaminmangel bedingen [17]. Das Symptomenspektrum umfasst bei allen Betroffenen geistige und körperliche Retardierung sowie ophthalmologische Auswirkungen [17]. Der Nachweis der Erkrankung erfolgt über eine genetische Analyse [17]. Die Therapie besteht in der Gabe von hoch dosiertem Vitamin B12 (parenteral), Biotin und Folsäure [17].

Zur Diagnostik von Vitamin-B12-Mangel, HU, PA und auch MMA kann auf Filterpapier aufgetropftes Trockenblut verwendet werden [22,23]. Für Vitamin-B12-Mangel, MMA und PA wird dieses unter anderem auf den Metabolit Propioncarnitin untersucht. Der Metabolit Methionin ist für die Erkennung einer HU nötig [22]. Beim in Deutschland gemäß der Kinder-Richtlinie des Gemeinsamen Bundesausschusses (G-BA) [24] durchgeführten erweiterten Neugeborenencreening (ENS) wird in der 36. bis 72. Lebensstunde Venen- oder Fersenblut gewonnen, auf Filterpapierkarten aufgetropft und hinsichtlich anderer Zielerkrankungen untersucht. Vitamin-B12-Mangel, HU, PA und MMA gehören bislang nicht zu den Zielerkrankungen, auf die im Rahmen des ENS untersucht wird.

Ziel eines Neugeborenencreenings auf Vitamin-B12-Mangel, HU, PA oder MMA ist die frühere Identifikation und Behandlung von betroffenen Kindern, bevor es zu einer metabolischen Entgleisung mit irreversiblen neurologischen Schäden im Sinne eines Residualsyndroms kommt.

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung einer Früherkennung eines Vitamin-B12-Mangels und weiterer Zielerkrankungen (Homocystinurie [HU], Propionazidämie [PA] und Methylmalonazidurie [MMA]) im erweiterten Neugeborenencreening (ENS)¹. Dabei wird das Neugeborenencreening auf die vier Zielerkrankungen im Vergleich zu keinem Screening hinsichtlich patientenrelevanter Endpunkte bewertet.

¹ Sofern nicht explizit anders vermerkt, wird der Begriff Zielerkrankungen nachfolgend synonym für Vitamin-B12-Mangel, HU, PA und MMA verwendet.

3 Projektverlauf

Der Gemeinsame Bundesausschuss (G-BA) hat am 24.11.2022 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung einer Früherkennung eines Vitamin B12-Mangels und weiterer Zielerkrankungen (Homocystinurie, Propionazidämie und Methylmalonazidurie) im erweiterten Neugeborenen-Screening beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Am 06.12.2022 wurden Betroffene konsultiert, um im Rahmen der Projektbearbeitung einen Eindruck davon zu gewinnen, wie Patientinnen und Patienten (oder ihre Angehörigen) die Erkrankungen erleben, welche Therapieerfahrungen sie gemacht haben und was sie sich von einer Behandlung erhoffen.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 4 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation der Anhörung zum Vorbericht veröffentlicht.

Der Zeitplan für alle Arbeitsschritte der Berichterstellung für dieses Projekt ist auf der Website des IQWiG dargelegt.

4 Methoden

Diese Bewertung wird auf Grundlage der Allgemeinen Methoden 7.0 [25] erstellt.

Der Nutzen eines Neugeborenen Screenings auf die Zielerkrankungen kann auf 2 Wegen bewertet werden. Diese Herangehensweisen werden im Folgenden beschrieben.

Nutzenbewertung anhand von vergleichenden Interventionsstudien der Screeningkette

Der Nutzen von Screeningmaßnahmen lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien mit einer (idealerweise randomisierten) Zuteilung von Personen zu einer Strategie mit beziehungsweise ohne Anwendung der Screeningmaßnahme wie auch anhand von vergleichenden Kohortenstudien der gesamten Screeningkette und der Betrachtung patientenrelevanter Endpunkte bewerten [25]. In solchen Studien erhalten die Neugeborenen in der Kontrollgruppe kein Screening auf die Zielerkrankungen. Die der Interventionsgruppe zugeteilten Neugeborenen erhalten ein Screening auf die Zielerkrankungen und ihnen werden entsprechend dem Untersuchungsergebnis des Screenings gegebenenfalls eine diagnostische Abklärung und Therapie zugewiesen.

Nutzenbewertung anhand von vergleichenden Interventionsstudien 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 wird der Nutzen eines früheren gegenüber einem späteren Therapiebeginn erfasst, die diagnostische Güte untersucht und gesundheitsbezogene Konsequenzen für falsch-positive, richtig-positive, falsch-negative sowie richtig-negative Befunde gegenübergestellt. Der Nutzen des Screenings kann dadurch abgeleitet werden, dass ein früherer gegenüber einem späteren Therapiebeginn einen Nutzen zeigt und gleichzeitig der Screeningtest eine hinreichende diagnostische Güte aufweist (Linked-Evidence-Ansatz). Dazu werden vergleichende Interventionsstudien zum Therapiebeginn und Studien zur diagnostischen Güte herangezogen.

4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette

4.1.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet das Neugeborenen Screening auf wenigstens eine der Zielerkrankungen. Der Zeitpunkt der Probenentnahme soll auf den in der Kinder-Richtlinie des

G-BA genannten Zeitrahmen für das ENS übertragbar sein (siehe Kapitel 1). Als Vergleichsintervention gilt kein Screening auf die Zielerkrankungen.

Die in den Studien zur Screeningkette angewendeten Diagnose- und Therapieverfahren müssen auf die Situation des Neugeborenen Screenings im Rahmen der Kinder-Richtlinie in Deutschland übertragbar sein.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtüberleben, krankheitsspezifisches Überleben),
- Morbidität (zum Beispiel Entwicklungsstörungen und Wachstumsstörungen, Krankenhausaufenthalte),
- (schwerwiegende) unerwünschte Ereignisse,
- gesundheitsbezogene Lebensqualität des Kindes.

4.1.4 Studientypen

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

Für den zu erstellenden Bericht werden in erster Linie RCTs als relevante wissenschaftliche Informationsquellen in die Nutzenbewertung einfließen.

Falls keine RCTs in ausreichender Zahl und / oder Qualität vorliegen, werden schrittweise auch Studien einer niedrigeren Evidenzstufe eingeschlossen.

1. RCTs,
2. quasirandomisierte kontrollierte Studien,
3. prospektive vergleichende Kohortenstudien.

Dabei erfolgt der Einschluss von nicht randomisierten vergleichenden Studien ausschließlich bei adäquater Confounderkontrolle. Eine adäquate Confounderkontrolle liegt vor, wenn das Problem einer möglichen Strukturungleichheit bei der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Hierfür müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können. Hierfür sind mindestens Daten zum Confounder Alter erforderlich.

Falls weder RCTs noch quasirandomisierte kontrollierte Studien oder prospektiv vergleichende Kohortenstudien in ausreichender Zahl und / oder Qualität vorliegen und dramatische Effekte möglich sind, werden die folgenden Studientypen schrittweise eingeschlossen:

4. retrospektive vergleichende Kohortenstudien mit zeitlich paralleler Kontrollgruppe²,
5. retrospektive vergleichende Kohortenstudien mit nicht zeitlich paralleler Kontrollgruppe².

Auch zu diesen Studien müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

² Im Rahmen der vorliegenden Bewertung werden hierunter auch Publikationen von Registerauswertungen gefasst, in denen 2 Kollektive miteinander verglichen werden und in denen die Auswertung prospektiv geplant worden ist.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)

Einschlusskriterien	
INS1	Neugeborene
INS2	Prüfintervention: Neugeborenencreening auf wenigstens 1 der Zielerkrankungen (Vitamin-B12-Mangel, Homocystinurie, Propionazidämie und Methylmalonazidurie) unter Verwendung von Filterpapierkarten (siehe auch Abschnitt 4.1.2)
INS3	Vergleichsintervention: kein Neugeborenencreening auf wenigstens 1 der Zielerkrankungen (siehe auch Abschnitt 4.1.2)
INS4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
INS5	Studientypen: RCTs, nicht randomisierte vergleichende Interventionsstudien, vergleichende Kohortenstudien (auch retrospektiv oder mit historischem Vergleich) (siehe auch Abschnitt 4.1.4)
INS6	Publikationssprache: Deutsch oder Englisch
INS7	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [26] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [27], TREND- [28] oder STROBE-Statements [29] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; INS: Interventionsstudien der Screeningkette; RCT: Randomized controlled Trial; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

4.2 Kriterien für den Einschluss von vergleichenden Interventionsstudien zu Therapiebeginn

Studien, die einen Vergleich eines früheren versus eines späteren Therapiebeginns ermöglichen, werden im Rahmen des vorliegenden Berichts systematisch recherchiert und ausgewertet, wenn vergleichende Interventionsstudien der Screeningkette nicht oder in nicht ausreichender Quantität und Qualität vorliegen.

4.2.1 Population

In die Bewertung werden jeweils Studien mit Patientinnen und Patienten mit HU, PA oder MMA bis zum Alter von 12 Jahren aufgenommen. Diese Festlegung erfolgt insbesondere daher, dass sich diese Erkrankungen auch im Kindesalter erst manifestieren können. Zudem werden somit solche Patientinnen und Patienten berücksichtigt, bei denen vergleichbare Auswirkungen der Zielerkrankungen zu erwarten sind. Für die Zielerkrankung Vitamin-B12-Mangel werden Studien mit Patientinnen und Patienten bis zum Alter von 12 Monaten aufgenommen. Diese Differenzierung erlaubt einen Fokus auf Vitamin-B12-Mangel wie er hinsichtlich Ursachen und Auswirkungen in Kapitel 1 beschrieben ist.

4.2.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet ein früherer Therapiebeginn. Als Vergleichsintervention gilt ein späterer Therapiebeginn. Die Therapie besteht in einer jeweils von der Zielerkrankung abhängigen spezifischen medikamentösen Behandlung sowie gegebenenfalls einer auf die zugrunde liegende Zielerkrankung abgestimmten Diät und gegebenenfalls weiterer unterstützender Maßnahmen.

Die in den Studien zum Therapiebeginn angewendeten Diagnose- und Therapieverfahren müssen auf die Situation des Neugeborenen Screenings / der -behandlung in Deutschland übertragbar sein.

4.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden die unter Abschnitt 4.1.3 genannten patientenrelevanten Endpunkte betrachtet.

4.2.4 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 einer medizinischen Intervention.

Für den zu erstellenden Bericht werden in erster Linie RCTs als relevante wissenschaftliche Informationsquelle in die Nutzenbewertung einfließen.

Falls keine RCTs in ausreichender Zahl und / oder Qualität vorliegen, werden schrittweise auch Studien einer niedrigeren Evidenzstufe eingeschlossen.

1. RCTs,
2. quasirandomisierte kontrollierte Studien,
3. prospektive vergleichende Kohortenstudien.

Dabei erfolgt der Einschluss von nicht randomisierten vergleichenden Studien ausschließlich bei adäquater Confounderkontrolle. Eine adäquate Confounderkontrolle liegt vor, wenn das Problem einer möglichen Strukturungleichheit bei der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Hierfür müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können.

Falls weder RCTs noch quasirandomisierte kontrollierte Studien oder prospektiv vergleichende Kohortenstudien in ausreichender Zahl und / oder Qualität vorliegen und

dramatische Effekte möglich sind, werden die folgenden Studientypen schrittweise eingeschlossen:

4. retrospektive vergleichende Kohortenstudien mit zeitlich paralleler Kontrollgruppe³,
5. retrospektive vergleichende Kohortenstudien mit nicht zeitlich paralleler Kontrollgruppe³.

Auch zu diesen Studien müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können.

4.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.2.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.2.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 2 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

³ Im Rahmen der vorliegenden Bewertung werden hierunter auch Publikationen von Registerauswertungen gefasst, in denen 2 Kollektive miteinander verglichen werden und in denen die Auswertung prospektiv geplant worden ist.

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

Einschlusskriterien	
INT1	Patientinnen und Patienten mit Homocystinurie bzw. Propionazidämie bzw. Methylmalonazidurie bis 12 Jahre bzw. Patientinnen und Patienten mit Vitamin-B12-Mangel bis 12 Monate; siehe auch Abschnitt 4.2.1)
INT2	Prüfintervention: frühere Behandlung (siehe auch Abschnitt 4.2.2)
INT3	Vergleichsintervention: spätere Behandlung (siehe auch Abschnitt 4.2.2)
INT4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
INT5	Studientypen: RCTs, nicht randomisierte vergleichende Interventionsstudien, vergleichende Kohortenstudien (auch retrospektiv oder mit historischem Vergleich) (siehe auch Abschnitt 4.2.4)
INT6	Publikationssprache: Deutsch oder Englisch
INT7	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [26] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [27], TREND- [28] oder STROBE-Statements [29] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; INT: Interventionsstudien zum Therapiebeginn; RCT: Randomized controlled Trial; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

4.3 Kriterien für den Einschluss von Studien zur diagnostischen Güte

Sollte sich aus den Interventionsstudien der Screeningkette oder den Interventionsstudien zum Therapiebeginn eine positive Aussage zum Nutzen eines früheren Therapiebeginns ergeben, werden im Rahmen des vorliegenden Berichts auch Studien zur diagnostischen Güte zur Nutzenbewertung herangezogen. Die Charakteristika und Einschlusskriterien für die Studien werden im Folgenden beschrieben.

4.3.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

4.3.2 Indextest

Als Indextest betrachtet werden alle in den Studien verwendeten diagnostischen Testverfahren oder Kombinationen von Testverfahren zur Testung auf mindestens eine der Zielerkrankungen unter Verwendung von Trockenblut auf Filterpapierkarten. Der Zeitpunkt der Probenentnahme soll auf den in der Kinder-Richtlinie des G-BA genannten Zeitrahmen für das ENS übertragbar sein. Die laboranalytische Methodik und die dazugehörigen Spezifikationen für den Test zur Unterscheidung positiver und negativer Ergebnisse müssen prospektiv festgelegt worden sein. Die Testentwicklung und -validierung müssen an voneinander unabhängigen Stichproben durchgeführt worden sein.

4.3.3 Referenztest

Für die Erkennung eines Vitamin-B12-Mangels werden solche Nachweise als Referenztest akzeptiert, die eine Bestimmung des Vitamin-B12-Serumspiegels sowie funktioneller Marker im Urin beinhalten. Für die weiteren Zielerkrankungen sind die Referenztests Nachweise des jeweils zugrunde liegenden Enzymdefekts durch genetische Analysen. Bei unauffälligem Befund im Indextest kann alternativ auch die Nachbeobachtung akzeptiert werden.

4.3.4 Zielgrößen

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

4.3.5 Studientypen

Um die diagnostische Güte des Indextests zur Erkennung der Zielerkrankungen bei Neugeborenen möglichst unverzerrt bestimmen zu können, soll eine Gruppe von Neugeborenen, die zu einem bestimmten Zeitpunkt prospektiv rekrutiert und gescreent wurde, zeitnah (Querschnittsdesign) mit dem krankheitsspezifischen Referenzstandard (nach)untersucht beziehungsweise bei unauffälligem Befund im Indextest nachbeobachtet werden (siehe Abschnitt 4.3.3). Dabei sind ein konsekutiver, das heißt nicht selektiver 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 der Zielerkrankungen auch diagnostische retrospektive Kohortenstudien aufgenommen.

Ist die Datenlage aus Studien, die sowohl positive als auch negative Ergebnisse im Indextest mit dem Referenztest direkt überprüfen (komplette Verifikation), unzureichend, können Studien im Verification-of-only-positive-Testers(VOPT)-Design herangezogen werden. Dabei werden alle positiven Ergebnisse im Indextest mit dem Referenztest untersucht [30] und es können Aussagen zum positiven prädiktiven Wert (PPV) als Maß der diagnostischen Güte getroffen werden. Eine Bewertung der testnegativen Fälle und damit eine Bestimmung der Sensitivität oder Spezifität des Tests ist mit solchen Studien nicht möglich.

4.3.6 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.3.7 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.3.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

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 (Studien zur diagnostischen Güte)

Einschlusskriterien	
D1	Neugeborene (siehe auch Abschnitt 4.3.1)
D2	Indextest: Screening auf wenigstens 1 der Zielerkrankungen (Vitamin-B12-Mangel, Homocystinurie, Propionazidämie und Methylmalonazidurie) unter Verwendung von Filterpapierkarten (siehe auch Abschnitt 4.3.2)
D3	Referenztest: beinhaltet eine Bestimmung des Vitamin-B12-Serumspiegels sowie funktioneller Marker bzw. einen Nachweis des jeweils zugrunde liegenden Enzymdefekts und /oder genetische Analysen; Nachbeobachtung (siehe auch Abschnitt 4.3.3)
D4	Zielgrößen: personenbezogene Daten zur diagnostischen Güte (siehe auch Abschnitt 4.3.4)
D5	Studientypen: diagnostische Querschnitt- und Kohortenstudien (siehe auch Abschnitt 4.3.5)
D6	Publikationssprache: Deutsch oder Englisch
D7	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des STARD- [31] oder STROBE-Statements [29] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology</p>	

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

Für die Einschlusskriterien INS1, INT1 und D1 (Population), INS2, INT2 und D2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie, beziehungsweise Indextest bei Diagnosestudien) und INS3 und INT3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen und Patienten vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien INS1, INT1 beziehungsweise D1, INS2, INT2 beziehungsweise D2 und INS3 beziehungsweise INT3 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

4.5 Informationsbeschaffung

4.5.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung des Berichtsplans erfolgte eine Recherche nach systematischen Übersichten in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), in der International HTA Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

Die Suche fand am 23.11.2022 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Es wird geprüft, ob mindestens eine hochwertige und aktuelle systematische Übersicht (SÜ) infrage kommt, deren Informationsbeschaffung als Grundlage für die Bewertung verwendet werden kann (im Folgenden: Basis-SÜ). Dafür erfolgt eine Bewertung der Qualität der Informationsbeschaffung dieser SÜ(s). Kann mindestens 1 diesbezüglich hochwertige und aktuelle Basis-SÜ identifiziert werden, werden die zugrunde liegenden Studien beziehungsweise Dokumente von 1 Person auf ihre Relevanz für die vorliegende Bewertung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der eingeschlossenen Studien oder die Datenextraktion werden nicht übernommen.

Die finale Entscheidung, ob und wenn ja welche SÜ(s) als Basis-SÜ(s) herangezogen wird / werden, erfolgt nach Fertigstellung des Berichtsplans anhand der darin festgelegten Kriterien. In jedem Fall werden die Referenzlisten der identifizierten SÜs hinsichtlich relevanter Primärstudien gesichtet (siehe 4.5.2).

4.5.2 Umfassende Informationsbeschaffung von Studien

Für die umfassende Informationsbeschaffung wird eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten durchgeführt.

Für den Fall, dass mindestens 1 systematische Übersicht als Basis-SÜ für die Informationsbeschaffung verwendet werden kann (siehe Abschnitt 4.5.1), wird diese für die Informationsbeschaffung von Studien für den von der Übersicht abgedeckten Zeitraum herangezogen. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den nicht von der Übersicht abgedeckten Zeitraum.

Für den Fall, dass keine Basis-SÜ identifiziert werden kann, findet eine systematische Recherche für den gesamten relevanten Zeitraum statt.

Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Embase
 - Cochrane Central Register of Controlled Trials
- Studienregister
 - U.S. National Institutes of Health; ClinicalTrials.gov
 - World Health Organization; International Clinical Trials Registry Platform Search Portal
 - European Medicines Agency; EU Clinical Trials Register (für die Suche nach vergleichenden Interventionsstudien zum Therapiebeginn)

Aufgrund der separaten Suchen in ClinicalTrials.gov werden aus dem Suchergebnis des International Clinical Trials Registry Platform Search Portal Einträge dieses Registers entfernt.

Weitere Informationsquellen und Suchtechniken

- durch den G-BA übermittelte Dokumente
- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter SÜs
- Anhörung zum Vorbericht
- Autorenanfragen

4.5.3 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von systematischen Übersichten

Die MEDLINE-Suchstrategie enthält Limitierungen auf deutsch- und englischsprachige Publikationen [25] sowie auf Humanstudien.

Umfassende Informationsbeschaffung von Studien

Prospektiv ist keine zeitliche Einschränkung vorgesehen. Sollte die Informationsbeschaffung auf Grundlage einer Basis-SÜ erfolgen, wird eine entsprechende zeitliche Einschränkung in Betracht gezogen (siehe Abschnitt 4.5.1). Mit den Suchstrategien werden folgende Publikationstypen ausgeschlossen: Kommentare (MEDLINE) und Editorials (MEDLINE, Embase), da diese i. d. R. keine Studien enthalten [32] sowie Conference Abstracts und

Conference Reviews (Embase) [25]. Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen [25] sowie auf Humanstudien (MEDLINE, Embase). In der Embase-Suche werden MEDLINE-Datensätze und in der Cochrane-Central Register-of-Controlled-Trials-Suche Einträge aus Studienregistern ausgeschlossen.

4.5.4 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Datenbanken

Duplikate werden mithilfe von EndNote entfernt. Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 1, Tabelle 2 und Tabelle 3) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Die Rechercheergebnisse aus den folgenden Informationsquellen werden von 2 Personen unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

- Studienregister,
- durch den G-BA übermittelte Dokumente.

Die Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person auf Studien gesichtet. Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden aufgelöst.

4.6 Informationsbewertung und -synthese

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert.

4.6.1 Darstellung der Einzelstudien

4.6.1.1 Darstellung der Einzelstudien (vergleichende Interventionsstudien)

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht beschrieben.

Die relevanten Ergebnisse werden endpunktspezifisch pro Studie auf ihr jeweiliges Verzerrungspotenzial überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in den Abschnitten 4.6.3 bis 4.6.5 beschriebenen Verfahren eingesetzt.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

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

4.6.1.2 Darstellung der Einzelstudien (Studien zur diagnostischen Güte)

Die Ergebnisse zu den in den Studien berichteten Zielgrößen werden im Bericht beschrieben.

Die relevanten Ergebnisse werden jeweils auf ihr Verzerrungspotenzial und ihre Übertragbarkeit überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in den Abschnitten 4.6.3 bis 4.6.5 beschriebenen Verfahren eingesetzt.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

4.6.2 Bewertung des Verzerrungspotenzials der Ergebnisse

4.6.2.1 Bewertung des Verzerrungspotenzials der Ergebnisse (vergleichende Interventionsstudien)

Das Verzerrungspotenzial der Ergebnisse wird endpunktspezifisch für jede in die Nutzenbewertung eingeschlossene Studie bewertet. Dazu werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Kriterien systematisch extrahiert und bewertet:

A: Kriterien für die endpunktübergreifende Bewertung des Verzerrungspotenzials der Ergebnisse

- 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 der Patientin oder des Patienten sowie der behandelnden Personen (bei randomisierten Studien)
- ergebnisunabhängige Berichterstattung

B: Kriterien für die endpunktspezifische Bewertung des Verzerrungspotenzials der Ergebnisse

- Verblindung der Endpunkterheber / Endpunkterheberinnen
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisunabhängige Berichterstattung

Für die Ergebnisse randomisierter Studien wird das Verzerrungspotenzial zusammenfassend als niedrig oder hoch eingestuft. Wird bereits hinsichtlich der unter (A) aufgeführten Kriterien ein endpunktübergreifend hohes Verzerrungspotenzial festgestellt, gilt dieses damit für alle Ergebnisse aller Endpunkte als hoch, unabhängig von der Bewertung endpunktspezifischer Aspekte. Andernfalls finden anschließend die unter (B) genannten Kriterien pro Endpunkt Berücksichtigung.

Das Verzerrungspotenzial der Ergebnisse nicht randomisierter vergleichender Studien wird aufgrund der fehlenden Randomisierung zusammenfassend grundsätzlich als hoch bewertet.

4.6.2.2 Bewertung des Verzerrungspotenzials der Ergebnisse (Studien zur diagnostischen Güte)

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der Studien zur diagnostischen Güte erfolgt auf Basis des Instruments QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) [33]. Das Verzerrungspotenzial von Studien zur diagnostischen Güte wird als niedrig oder hoch eingestuft.

4.6.3 Metaanalysen

4.6.3.1 Metaanalysen (vergleichende Interventionsstudien)

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend wird die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [34] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall

von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten gemäß der Methode nach Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [35]. Als Ergebnis wird der gemeinsame Effekt inklusive Konfidenzintervall dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Dazu müssen die Studien ausreichend ähnlich sein und es darf keine Gründe geben, die gegen die Anwendung eines Modells mit festem Effekt sprechen. Ist ein Modell mit festem Effekt nicht vertretbar, so wird wie folgt verfahren: Da das bevorzugte Verfahren nach Knapp und Hartung im Fall von sehr wenigen (< 5) Studien extrem breite und nicht informative Konfidenzintervalle liefern kann, wird die Schätzung nach diesem Verfahren im Fall von 2 Studien im Regelfall nicht dargestellt. Im Fall von 3 oder mehr Studien wird zunächst geprüft, ob das Konfidenzintervall des Knapp-Hartung-Verfahrens schmäler ist als das Konfidenzintervall mittels des DerSimonian-Laird-Verfahrens. Ist dies der Fall, wird die Schätzung nach Knapp und Hartung mit Ad-hoc-Varianzkorrektur weiter betrachtet. Des Weiteren wird geprüft, ob das 95 %-Konfidenzintervall des Verfahrens nach Knapp und Hartung (ggf. mit Ad-hoc-Varianzkorrektur) zu breit ist, z. B. wenn die Vereinigung der 95 %-Konfidenzintervalle der Einzelstudien enthalten ist. Ist dies nicht der Fall, wird das Knapp-Hartung-Konfidenzintervall zur Ableitung einer Nutzaussage herangezogen, wenn das Ergebnis statistisch signifikant ist. Wird die Schätzung mittels Knapp und Hartung (ggf. mit Ad-hoc-Varianzkorrektur) nicht herangezogen (regelmäßig im Fall von 2 Studien bzw. bei zu breitem 95 %-Konfidenzintervall) oder liegt kein statistisch signifikantes Ergebnis vor, so wird geprüft, ob das Verfahren nach DerSimonian und Laird ein statistisch signifikantes Ergebnis liefert. Ist dies der Fall, werden die Studienergebnisse qualitativ zusammengefasst. Ist das Ergebnis nach DerSimonian und Laird nicht statistisch signifikant, wird auf die qualitative Zusammenfassung verzichtet, da in diesem Fall auch ein korrektes Verfahren keinen statistisch signifikanten Effekt ergeben kann.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird im Fall von mindestens 5 Studien nur das Prädiktionsintervall dargestellt. Bei 4 oder weniger Studien erfolgt eine qualitative Zusammenfassung. In beiden Fällen wird außerdem untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt 4.6.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.6.5).

Abgesehen von den genannten Modellen können in bestimmten Situationen und mit besonderer Begründung Alternativen wie z. B. das Betabinomialmodell bei binären Daten [36] angewendet werden.

4.6.3.2 Metaanalysen (Studien zur diagnostischen Güte)

Die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle [37] aus den Studien werden mittels Forest Plot zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen erfüllt sind, für die Testgütekriterien eine Metaanalyse anhand der Sensitivität und Spezifität in einem bivariaten Modell durchgeführt [38]. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [39,40]. 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 Metanalyse der Regel nicht dargestellt.

Falls die bivariaten Metaanalysen präzise Schätzungen liefern, so werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität 2-dimensional grafisch dargestellt. Des Weiteren werden die aus der bivariaten Metanalyse gewonnenen Schätzungen für die Erwartungswerte als gepooltes Paar der Sensitivität und der Spezifität mit der dazugehörigen 95 %-Konfidenzregion dargestellt [41].

In Ausnahmefällen, wie beispielsweise beim Vorliegen von mehreren großen Studien mit niedrigerem Verzerrungspotenzial, werden die Ergebnisse geeigneter univariater statistischer Tests, das heißt für die Sensitivität und Spezifität getrennt, dargestellt.

Sollten die Sensitivität und Spezifität nicht berechenbar sein – zum Beispiel, weil nur Studien im VOPT-Design eingeschlossen werden –, wird der PPV dargestellt und metaanalytisch zusammengefasst. Hierbei kommen Likelihood-basierte Verfahren auf Basis der individuellen Patientendaten zum Einsatz. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [42].

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen untersucht.

4.6.4 Sensitivitätsanalysen

Bestehen Zweifel an der Robustheit von Ergebnissen wegen methodischer Faktoren, die beispielsweise durch die Wahl bestimmter Cut-off-Werte, Ersetzungsstrategien für fehlende Werte, Erhebungszeitpunkte oder Effektmaße begründet sein können, ist geplant, den Einfluss solcher Faktoren in Sensitivitätsanalysen zu untersuchen. Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestuft Effekt kann zum Beispiel dazu führen,

dass nur ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.6.6).

4.6.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Entsprechende Auswertungen werden sowohl für Effekte als auch für Testgüte-Ergebnisse erfolgen. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter,
- Zielerkrankung,
- Dauer zwischen Symptom- und Therapiebeginn,
- diagnostische Testverfahren.

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 (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen und Patienten eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.6.6).

4.6.6 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens oder (höheren) Schadens getroffen. Dabei sind 4 Abstufungen der Aussagesicherheit möglich: 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 (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die regelhaft abzuleitende Aussagesicherheit ist von den in Tabelle 4 dargestellten Kriterien abhängig. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie. 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.

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

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			gemeinsame Effektschätzung sinnvoll	gemeinsame Effektschätzung nicht sinnvoll		
				Metaanalyse statistisch signifikant	konkludente Effekte ^a	
			deutlich	mäßig	nein	
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–

a. Unter konkludenten Effekten wird eine Datensituation verstanden, in der es möglich ist, einen Effekt im Sinne der Fragestellung abzuleiten, obwohl eine gemeinsame Effektschätzung nicht sinnvoll möglich ist (siehe Abschnitt 3.1.4 der Allgemeinen Methoden [25]).

Abschließend erfolgt eine endpunktübergreifende Bewertung des Nutzens.

Der Nutzen des Screenings kann auch durch die Gegenüberstellung der gesundheitsbezogenen Konsequenzen der möglichen Testergebnisse und ihrer Wahrscheinlichkeiten zusammen mit einer Aussage zum Nutzen eines früheren Therapiebeginns abgeleitet werden. Die Aussagesicherheit bezüglich des Nutzens des Screenings berücksichtigt dann sowohl die Aussagesicherheit bezüglich des Nutzens eines früheren Therapiebeginns als auch das Verzerrungspotenzial bezüglich der diagnostischen Güte.

In die abschließende endpunktübergreifende Bewertung des Nutzens wird jeweils auch die Datenvollständigkeit und die sich daraus möglicherweise ergebende Verzerrung aufgrund von Publikationsbias einbezogen, die zur Einschränkung des Fazits führen kann.

5 Literatur

1. Gramer G, Hoffmann GF. Vitamin-B12-Mangel im Neugeborenen- und Säuglingsalter – Ursachen, Früherkennung, Diagnostik und Vorstellung eines primär oralen Behandlungsschemas. Monatsschr Kinderheilkd 2020; (9).
<https://dx.doi.org/10.1007/s00112-020-01008-5>.
2. Chao MM, Illsinger S, Yoshimi A et al. Congenital Transcobalamin II Deficiency: A Rare Entity with a Broad Differential. Klin Padiatr 2017; 229(6): 355-357.
<https://dx.doi.org/10.1055/s-0043-120266>.
3. Hawthorne S, Levy HL. Can Newborn Screening for Vitamin B12 Deficiency be Incorporated into All Newborn Screening Programs? J Pediatr 2020; 216(1): 9-11.e1.
<https://dx.doi.org/10.1016/j.jpeds.2019.08.061>.
4. orphanet. Transcobalamin II-Mangel [online]. 2013 [Zugriff: 11.01.2023]. URL:
https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=1729.
5. Ludwig Boltzmann Institut. Eltern-Kind-Vorsorge neu; Update Teil IX: Empfehlungen aus evidenzbasierten Leitlinien für Screenings von Kindern (0–6 Jahre); Endbericht; LBI-HTA-Projektbericht Nr. 62, Update 2016. [online]. 2016 [Zugriff: 21.12.2022]. URL:
https://eprints.aihta.at/996/5/HTA-Projektbericht_Nr.62_Update_2016.pdf.
6. Huemer M, Diodato D, Schwahn B et al. Guidelines for diagnosis and management of the cobalamin-related remethylation disorders cblC, cblD, cblE, cblF, cblG, cblJ and MTHFR deficiency. J Inherit Metab Dis 2017; 40(1): 21-48. <https://dx.doi.org/10.1007/s10545-016-9991-4>.
7. Huemer M, Kozich V, Rinaldo P et al. Newborn screening for homocystinurias and methylation disorders: systematic review and proposed guidelines. J Inherit Metab Dis 2015; 38(6): 1007-1019. <https://dx.doi.org/10.1007/s10545-015-9830-z>.
8. Walter JH, Jahnke N, Remington T. Newborn screening for homocystinuria. Cochrane Database Syst Rev 2015; (10): CD008840.
<https://dx.doi.org/10.1002/14651858.CD008840.pub4>.
9. orphanet. Homocystinurie, klassische [online]. 2007 [Zugriff: 21.12.2022]. URL:
https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=173.
10. Morris AA, Kozich V, Santra S et al. Guidelines for the diagnosis and management of cystathionine beta-synthase deficiency. J Inherit Metab Dis 2017; 40(1): 49-74.
<https://dx.doi.org/10.1007/s10545-016-9979-0>.
11. orphanet. Homocystinurie ohne Methylmalonazidurie [online]. 2012 [Zugriff: 21.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=DE&Expert=622.

12. Diekman EF, de Koning TJ, Verhoeven-Duif NM et al. Survival and psychomotor development with early betaine treatment in patients with severe methylenetetrahydrofolate reductase deficiency. JAMA Neurology 2014; 71(2): 188-194. <https://dx.doi.org/10.1001/jamaneurol.2013.4915>.
13. Forny P, Hörster F, Ballhausen D et al. Guidelines for the diagnosis and management of methylmalonic acidemia and propionic acidemia: First revision. J Inher Metab Dis 2021; 44(3): 566-592. <https://dx.doi.org/10.1002/jimd.12370>.
14. Van Gosen L. Organic acidemias: a methylmalonic and propionic focus. J Pediatr Nurs 2008; 23(3): 225-233. <https://dx.doi.org/10.1016/j.pedn.2008.02.004>.
15. orphanet. Propionazidämie [online]. 2014 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=3557.
16. orphanet. Methylmalonazidämie ohne Homocystinurie [online]. 2012 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=20854.
17. orphanet. Methylmalonazidämie mit Homocystinurie [online]. 2013 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=DE&Expert=26.
18. orphanet. Methylmalonazidämie, Vitamin B12-sensible [online]. 2012 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=3260.
19. orphanet. Methylmalonazidämie, Vitamin B12-resistente, Typ mut0 [online]. 2012 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=20777.
20. orphanet. Methylmalonazidämie, Vitamin B12-resistente, Typ mut- [online]. 2012 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=11338.
21. orphanet. Methylmalonazidämie, Vitamin B12-resistente [online]. 2012 [Zugriff: 22.12.2022]. URL: https://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=DE&data_id=1263.
22. Gramer G, Fang-Hoffmann J, Feyh P et al. Newborn Screening for Vitamin B12 Deficiency in Germany - Strategies, Results, and Public Health Implications. J Pediatr 2020; 216: 165-172.e4. <https://dx.doi.org/https://doi.org/10.1016/j.jpeds.2019.07.052>.
23. Gemeinsamer Bundesausschuss. Unterausschuss Methodenbewertung (UA MB);

Antrag der Patientenvertretung nach § 140f SGB V auf Bewertung eines Neugeborenen-Screenings gemäß § 135 Abs. 1 i. V. m. § 26 SGB V; Früherkennung eines Vitamin B12-Mangels und weiterer Zielerkrankungen (Homocystinurie, Propionazidämie und Methylmalonazidurie) im erweiterten Neugeborenen-Screening (ENS) [online]. 2022 [Zugriff: 22.12.2021]. URL: https://www.g-ba.de/downloads/40-268-8937/2022-10-20_Einl-Beratung_Vitamin-B12-Mangel-erw-Neugeborenen-Screening_Antrag-PatV.pdf.

24. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über die Früherkennung von Krankheiten bei Kindern (Kinder-Richtlinie) [online]. 2022 [Zugriff: 07.10.2022]. URL: https://www.g-ba.de/downloads/62-492-2848/Kinder-RL_2022-04-21_iK-2022-06-23.pdf.

25. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Entwurf für Version 7.0 [online]. 2022 [Zugriff: 06.12.2022]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_entwurf-fuer-version-7.pdf.

26. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 25.11.2020]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.

27. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. <https://dx.doi.org/10.1136/bmj.c869>.

28. Des Jarlais DC, Lyles C, Crepaz N et al. 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. <https://dx.doi.org/10.2105/ajph.94.3.361>.

29. Von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577. <https://dx.doi.org/10.7326/0003-4819-147-8-200710160-00010>.

30. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. Biostatistics 2001; 2(3): 249-260. <https://dx.doi.org/10.1093/biostatistics/2.3.249>.

31. Bossuyt PM, Reitsma JB, Bruns DE et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003; 138(1): 40-44. <https://dx.doi.org/10.7326/0003-4819-138-1-200301070-00010>.

32. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. <https://dx.doi.org/10.1002/jrsm.1425>.

33. Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-536. <https://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009>.
34. Sutton AJ, Abrams KR, Jones DR et al. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
35. Veroniki AA, Jackson D, Viechtbauer W et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015; (Suppl 1): 25-27.
36. Kuss O. Statistical methods for meta-analyses including information from studies without any events-add nothing to nothing and succeed nevertheless. *Stat Med* 2015; 34(7): 1097-1116. <https://dx.doi.org/10.1002/sim.6383>.
37. Leemis LM, Trivedi KS. A Comparison of Approximate Interval Estimators for the Bernoulli Parameter. *The American Statistician* 1996; 50(1): 63-68. <https://dx.doi.org/10.1080/00031305.1996.10473544>.
38. Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990. [https://dx.doi.org/S0895-4356\(05\)00162-9](https://dx.doi.org/S0895-4356(05)00162-9) [pii] [10.1016/j.jclinepi.2005.02.022](https://doi.org/10.1016/j.jclinepi.2005.02.022).
39. 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; author reply 1332-1333. <https://dx.doi.org/10.1016/j.jclinepi.2006.06.011>.
40. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-62, 62-54. <https://dx.doi.org/10.3414/ME09-01-0001>.
41. Hotelling H. The Generalization of Student's Ratio. *The Annals of Mathematical Statistics* 1931; 2(3): 360-378, 319.
42. Trikalinos TA, Trow P, Schmid CH. Simulation-Based Comparison of Methods for Meta-Analysis of Proportions and Rates; *Methods Research Report* : 13(14)-EHC084-EF [online]. 2013 [Zugriff: 28.12.2022]. URL: <https://www.ncbi.nlm.nih.gov/pubmed/24404633>.
43. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc* 2006; 94(4): 451-455.

Anhang A Suchstrategien in bibliografischen Datenbanken

A.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to November 22, 2022

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [43] – High specificity strategy

#	Searches
1	exp Vitamin B12/ or exp Vitamin B 12 Deficiency/
2	(vitamin* and (b adj1 "12")).ti,ab.
3	((vitamin* and b12) or (cobalamin* adj6 deficien*)).ti,ab.
4	or/1-3
5	exp infant/
6	(newborn* or neonat* or pediatric* or infant*).ti,ab.
7	or/5-6
8	and/4,7
9	(homocystinuria* or ((cystathionine* or mthfr*) adj6 deficien*)).mp.
10	((methylmalonic or malonic) adj6 (acid?emia* or aciduria*)).mp.
11	((propionic adj6 (acid?emia* or aciduria*)) or propionicaciduria* or propionicacidemia* or (propionyl adj1 CoA adj6 carboxylase adj6 deficien*) or (ketotic adj6 (glycinemia* or hyperglycinemia*))).mp.
12	or/8-11
13	Cochrane database of systematic reviews.jn.
14	(search or MEDLINE or systematic review).tw.
15	meta analysis.pt.
16	or/13-15
17	16 not (exp animals/ not humans.sh.)
18	and/12,17
19	18 and (english or german or multilingual or undetermined).lg.

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	"Vitamin B 12"[mhe]
2	"Vitamin B 12 Deficiency"[mhe]
3	((vitamin* AND (b12 OR b 12)) OR (cobalamin* AND deficien*)) [Title] OR ((vitamin* AND (b12 OR b 12)) OR (cobalamin* AND deficien*)) [abs]
4	#3 OR #2 OR #1
5	"Infant"[mh]
6	(newborn* OR neonat* OR pediatric* OR infant*) [Title] OR (newborn* OR neonat* OR pediatric* OR infant*) [abs]
7	#6 OR #5
8	#7 AND #4
9	(homocystinuria* OR ((cystathionine* OR mthfr*) AND deficien*)) [Title] OR (homocystinuria* OR ((cystathionine* OR mthfr*) AND deficien*)) [abs]
10	((methylmalonic OR malonic) AND (acidemia* OR acidaemia* OR aciduria*)) [Title] OR ((methylmalonic OR malonic) AND (acidemia* OR acidaemia* OR aciduria*)) [abs]
11	((propionic AND (acidaemia* OR acidemia* OR aciduria*)) OR propionicaciduria* OR propionicacidemia* OR (propionyl AND carboxylase AND deficien*) OR (ketotic AND (glycinemia* OR hyperglycinemia*))) [Title] OR ((propionic AND (acidaemia* OR acidemia* OR aciduria*)) OR propionicaciduria* OR propionicacidemia* OR (propionyl AND carboxylase AND deficien*) OR (ketotic AND (glycinemia* OR hyperglycinemia*))) [abs]
12	#11 OR #10 OR #9 OR #8