

**Nicht invasive Bestimmung
des fetalen Rhesusfaktors zur
Vermeidung einer
mütterlichen Rhesus-
sensibilisierung**

Vorläufiger Berichtsplan

Auftrag: D16-01
Version: 1.0
Stand: 11.01.2017

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Nutzenbewertung einer nicht invasiven Bestimmung des fetalen Rhesusfaktors zur Vermeidung einer mütterlichen Rhesussensibilisierung

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

22.09.2016

Interne Auftragsnummer:

D16-01

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

Bei dem vorliegenden Berichtsplan handelt es sich um eine vorläufige Version. Zu diesem Berichtsplan können alle interessierten Personen, Institutionen und Gesellschaften Stellungnahmen abgeben, die zu einer Ergänzung und / oder Überarbeitung des Berichtsplans führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG (www.iqwig.de), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

Schlagwörter: Genotypisierungstechniken, Rh-Hr-Blutgruppensystem, Fetus, Nutzenbewertung, Systematische Übersicht

Keywords: Genotyping Techniques, Rh-Hr Blood-Group System, Fetus, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abbildungsverzeichnis	vi
Abkürzungsverzeichnis	vii
1 Hintergrund	1
2 Fragestellung	3
3 Projektverlauf	4
3.1 Zeitlicher Verlauf des Projekts	4
4 Methoden	5
4.1 Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Untersuchung	8
4.1.1 Population.....	8
4.1.2 Prüf- und Vergleichsintervention	8
4.1.3 Patientenrelevante Endpunkte	8
4.1.4 Studientypen	9
4.1.5 Studiendauer	9
4.1.6 Tabellarische Darstellung der Kriterien für den Studieneinschluss	9
4.1.7 Studien im Anreicherungsdesign.....	10
4.2 Kriterien für den Einschluss von Studien zum Unterlassen einer Anti-D- Prophylaxe in die Untersuchung	10
4.2.1 Population.....	10
4.2.2 Prüf- und Vergleichsintervention	11
4.2.3 Patientenrelevante Endpunkte	11
4.2.4 Studientypen	11
4.2.5 Studiendauer	11
4.2.6 Tabellarische Darstellung der Kriterien für den Studieneinschluss	12
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	12
4.3.4 Zielgrößen	12
4.3.5 Studientypen	13
4.3.6 Studiendauer	13
4.3.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	13

4.4	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	13
4.5	Informationsbeschaffung.....	14
4.5.1	Primäre Suchquellen.....	14
4.5.1.1	Bibliografische Recherche.....	14
4.5.1.2	Öffentlich zugängliche Studienregister	14
4.5.2	Weitere Suchquellen.....	14
4.5.2.1	Systematische Übersichten	14
4.5.2.2	Öffentlich zugängliche Dokumente von Zulassungsbehörden	14
4.5.2.3	Durch den G-BA übermittelte Dokumente.....	15
4.5.2.4	Anhörung	15
4.5.2.5	Autorenanfragen	15
4.5.3	Selektion relevanter Studien.....	15
4.6	Informationsbewertung.....	16
4.6.1	Bewertung von vergleichenden Interventionsstudien	16
4.6.2	Bewertung von Studien zur diagnostischen Güte.....	17
4.7	Informationssynthese und -analyse	17
4.7.1	Gegenüberstellung der Ergebnisse der Einzelstudien	18
4.7.2	Meta-Analysen	18
4.7.2.1	Meta-Analysen für vergleichende Interventionsstudien.....	18
4.7.2.2	Meta-Analysen für Studien zur diagnostischen Güte	19
4.7.3	Aussagen zur Beleglage	19
4.7.4	Sensitivitätsanalysen	20
4.7.5	Subgruppenmerkmale und andere Effektmodifikatoren	21
5	Literatur	22

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostisch-therapeutischen Behandlungskette)	10
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (Studien zum Unterlassen nicht indizierter Anti-D-Prophylaxe)	12
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)	13
Tabelle 4: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit.....	20

Abbildungsverzeichnis

	Seite
Abbildung 1: Darstellung potenzieller Effekte der diagnostisch-therapeutischen Behandlungskette im Vergleich zwischen bisherigem (oben) und neuem Vorgehen (unten) unter Hinzunahme des Pränataltests	6

Abkürzungsverzeichnis

Abkürzung	Bedeutung
Anti-D-Ig	gegen Protein D des Rhesus-Blutgruppensystems gerichtetes Immunglobulin
DNA	Desoxyribonukleinsäure, Träger der Erbinformation
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
Rh(D)	Protein D des Rhesus-Blutgruppensystems

1 Hintergrund

Rhesus-D(Rh[D])-negative Schwangere können bei Übertritt von fetalem Blut in den mütterlichen Blutkreislauf Anti-D-Antikörper bilden, wenn das Kind Rh(D)-positiv ist. Diesen Vorgang bezeichnet man als Sensibilisierung und diese führt zu einer Inkompatibilität zwischen mütterlichem und fetalem Blut. Sie wird manifest in einer Folgeschwangerschaft mit einem abermals Rh(D)-positivem Kind und kann aufgrund der Plazentagängigkeit der IgG-Antikörper zu einem Abbau der fetalen Erythrozyten und in dessen Folge zu schwerwiegenden hämolytischen Erkrankungen des Kindes wie Anämie, Hydrops und Fruchttod führen. In 25 bis 35 % der Fälle von Rh(D)-Inkompatibilität kommt es bei normalem Schwangerschaftsverlauf bei der Geburt zu einer Anämie mit Bilirubinämie, die ohne Therapie zum Kernikterus und damit zu Hirnschäden führt. Eine neonatale Anämie wird mittels Fototherapie und gegebenenfalls mittels eines Blutaustauschs behandelt. Weitere 20 bis 25 % der Feten bei Rhesusinkompatibilität entwickeln bereits im Mutterleib eine hämolytische Anämie, wodurch es zu einem Hydrops fetalis mit Herzinsuffizienz und Fruchttod kommen kann. Beim erstmaligen Auftreten einer solchen Anämie ist / sind je nach Gestationsalter eine oder mehrere Bluttransfusionen in die Nabelschnurvene indiziert [1].

Laut den Mutterschafts-Richtlinien [2] werden zu einem möglichst frühen Zeitpunkt aus einer Blutprobe der Schwangeren die mütterliche Blutgruppe und der Rh(D)-Faktor bestimmt sowie ein Antikörper-Suchtest durchgeführt. Ein weiterer Antikörper-Suchtest ist bei allen Schwangeren in der 24. bis 27. Schwangerschaftswoche vorgesehen. Sind bei Rh(D)-negativen Schwangeren keine Anti-D-Antikörper nachweisbar, erhalten diese prophylaktisch in der 28. bis 30. Schwangerschaftswoche eine Standarddosis (300 µg) Anti-D-Immunglobulin, um eventuell vom fetalen in den mütterlichen Kreislauf übertretende Erythrozyten im Zeitraum bis zur Geburt abzufangen und damit eine Sensibilisierung zu verhindern. Bei jedem Kind einer Rh(D)-negativen Mutter ist unmittelbar nach der Geburt der Rh(D)-Faktor des Kindes zu bestimmen. Bei Rh(D)-positivem Kind ist der Mutter innerhalb von 72 Stunden post partum eine weitere Standarddosis Anti-D-Immunglobulin zu applizieren. Durch diese Prophylaxe soll ein schneller Abbau der insbesondere während der Geburt in den mütterlichen Kreislauf übergetretenen Rh(D)-positiven Erythrozyten bewirkt werden, um die Bildung von Rh(D)-Antikörpern bei der Mutter zu verhindern. Auch nach einer Fehlgeburt oder einem Schwangerschaftsabbruch [2] sowie nach einem Bauchtrauma oder anderen Vorkommnissen mit der Möglichkeit des Übertritts fetalen Bluts in den mütterlichen Kreislauf wird eine Anti-D-Prophylaxe gegeben [1]. Dabei handelt es sich um humanes Anti-D-Immunglobulin, das von sensibilisierten Spendern gewonnen wird. Als Nebenwirkungen der Präparate werden z. B. allergische Überempfindlichkeitsreaktionen (selten) und Hautreaktionen (gelegentlich) genannt. Die Möglichkeit der Übertragung von Erregern bei der Anwendung von aus menschlichem Blut oder Plasma hergestellten Arzneimitteln könne nicht vollständig ausgeschlossen werden [3,4].

Erhält die Mutter keine Prophylaxe, kommt es bei 4 bis 9 % der Geburten eines Rh(D)-positiven Kindes zu einer Sensibilisierung, ebenso in 4 % der Fälle nach einer Abortkürettage

oder vaginalen Blutung und in 2 bis 5 % der Fälle nach einer Chorionzottenbiopsie oder Amniozentese. Im 2. und 3. Trimenon kommt es spontan bei etwa 1 % zu einer Sensibilisierung [1].

Der Anteil Rh(D)-negativer Menschen in der mitteleuropäischen Gesamtbevölkerung beträgt 15 %. Bei rund 12,5 % aller europäischen Paare besteht eine Rhesuskonstellation während der Schwangerschaft, das heißt, die Mutter ist Rh(D)-negativ und der Vater Rh(D)-positiv. Da ein Teil der heterozygoten Väter das Rh(D)-Gen nicht weitervererbt, reduziert sich der Anteil Rh(D)-negativer Schwangerer mit Rh(D)-positiven Feten auf rund 8,2 % [1]. Da in Deutschland der Rh(D)-Status des Vaters nicht regelhaft bestimmt wird, erhalten insgesamt 6,8 % der Schwangeren bei dem derzeitigen Verfahren eine Anti-D-Prophylaxe, für die keine Indikation besteht. Bei 737 575 Geburten pro Jahr in 2015 [5] betrifft dies etwa 50 000 Schwangere in Deutschland.

Die zu prüfende Intervention stellt einen nicht invasiven molekulargenetischen Test dar, der anhand zellfreier fetaler DNA aus dem mütterlichen Plasma den Rh(D)-Faktor des Fetus auf Basis der Polymerase-Kettenreaktion (Polymerase Chain Reaction [PCR]) bestimmt. Erhält eine Rh(D)-negative Schwangere ein negatives Testergebnis und trägt damit ein Rh(D)-negatives Kind aus, könnte aufgrund des neuen Tests auf die Gabe von nicht indiziertem Anti-D-Immunglobulin verzichtet werden. Zudem ist es denkbar, dass durch Einführung des zusätzlichen Tests die Adhärenz hinsichtlich der Anti-D-Prophylaxe verbessert werden könnte, da den Frauen dadurch das Risiko bei zukünftigen Schwangerschaften bekannt wird. Dadurch könnte gegebenenfalls die Zahl von Sensibilisierungen gesenkt werden [6].

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung einer nicht invasiven Bestimmung des fetalen Rhesusfaktors bei Rh(D)-negativen Schwangeren zur Vermeidung einer mütterlichen Rhesussensibilisierung im Rahmen der Vorsorgeuntersuchungen gemäß den Mutterschafts-Richtlinien hinsichtlich patientenrelevanter Endpunkte.

Die Bewertung wird durchgeführt im Vergleich zur gegenwärtig angewandten Methode der regelhaften Gabe der Anti-D-Prophylaxe an alle Rh(D)-negativen Schwangeren ohne prä-natale Bestimmung des Rh(D)-Faktors des Fetus.

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat am 22.09.2016 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung einer nicht invasiven Bestimmung des fetalen Rhesusfaktors zur Vermeidung einer mütterlichen Rhesussensibilisierung beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans war eine Konsultation von Betroffenen unter anderem zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen vorgesehen. Trotz Anfragen bei verschiedenen Patientenorganisationen kam eine solche Konsultation nicht zustande.

Der vorliegende vorläufige Berichtsplan (Version 1.0) wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Stellungnahmen können alle interessierten Personen, Institutionen und Gesellschaften abgeben. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen. Diese sind ebenfalls auf der Website des IQWiG in einem Leitfaden dargelegt. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichtsplans führen. Im Anschluss an diese Anhörung wird der dann gültige Berichtsplan publiziert.

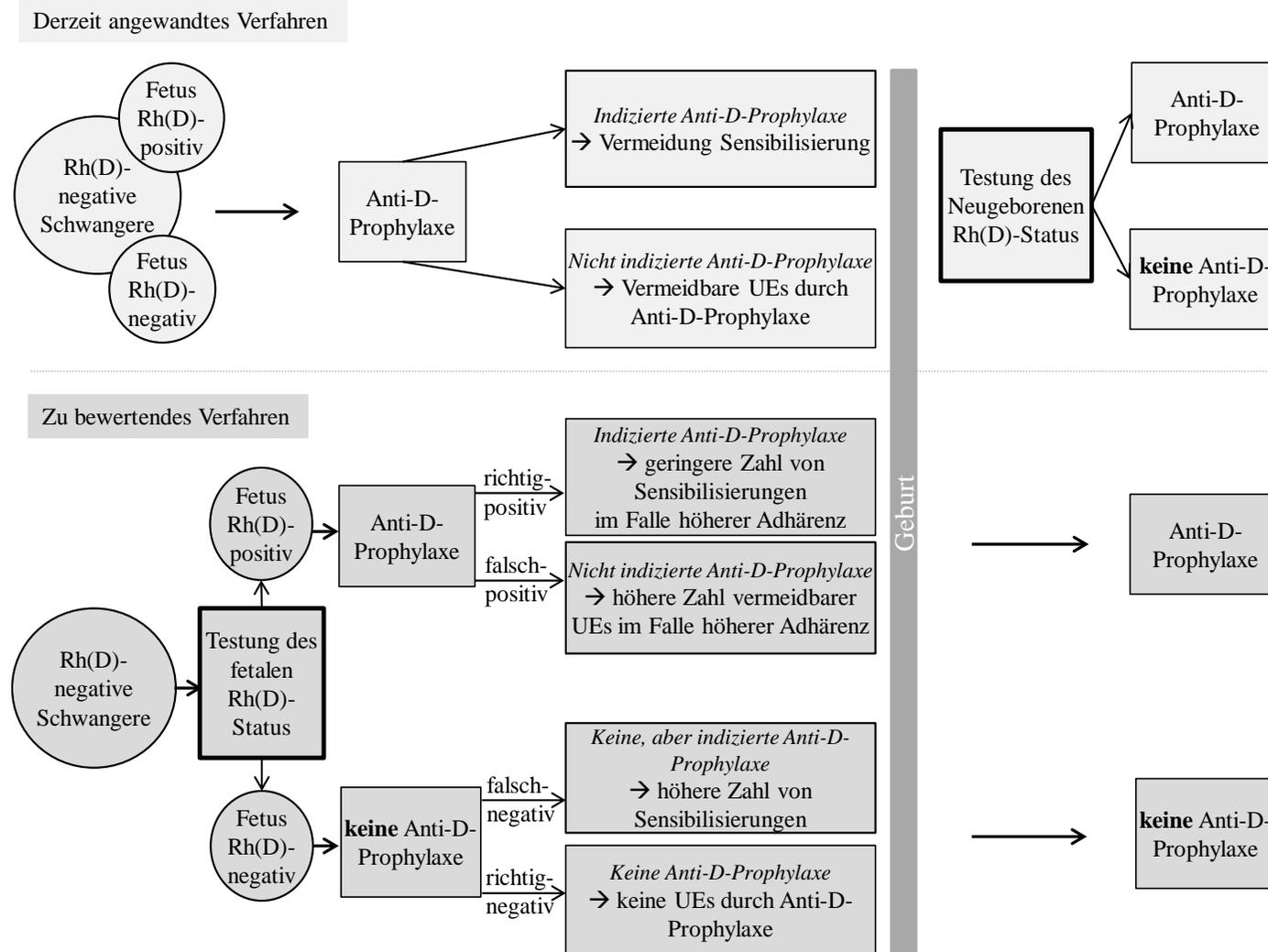
4 Methoden

Der Nutzen des pränatalen Tests zur Bestimmung des Rhesusfaktors des Fetus kann auf 2 Wegen bewertet werden. Diese Herangehensweisen werden im Folgenden beschrieben.

Nutzenbewertung anhand von vergleichenden Interventionsstudien der diagnostisch-therapeutischen Behandlungskette

Der Nutzen von diagnostischen Verfahren lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien der gesamten diagnostisch-therapeutischen Behandlungskette bewerten. Dabei werden Personen (idealerweise zufällig) unterschiedlichen Strategien zugeteilt [7]. Zur Bewertung werden patientenrelevante Endpunkte betrachtet.

Abbildung 1 stellt die potenziellen Effekte der diagnostisch-therapeutischen Behandlungskette unter Hinzunahme des zu bewertenden Pränataltests dar.



Rh(D): Protein D des Rhesus-Blutgruppensystems; UEs: unerwünschte Ereignisse

Abbildung 1: Darstellung potenzieller Effekte der diagnostisch-therapeutischen Behandlungskette im Vergleich zwischen bisherigem (oben) und neuem Vorgehen (unten) unter Hinzunahme des Pränataltests

In die Bewertung fließen auch Studien im Anreicherungsdesign ein, die ausschließlich Effekte bei Schwangeren mit einem bestimmten Testergebnis (Fetus Rh[D]-positiv oder Rh[D]-negativ) untersuchen.

Details zu diesem Vorgehen finden sich in Abschnitt 4.1.

Nutzenbewertung anhand von vergleichenden Interventionsstudien zur Therapie und Studien zur Bewertung der diagnostischen Güte

Liegen vergleichende Interventionsstudien der diagnostisch-therapeutischen Behandlungskette für die Nutzenbewertung nicht oder in nicht ausreichender Quantität und Qualität vor, kann eine Bewertung der einzelnen Bausteine der diagnostisch-therapeutischen Behandlungskette erfolgen. Für die Nutzenbewertung werden gesundheitsbezogene Konsequenzen für falsch-positive, richtig-positive, falsch-negative sowie richtig-negative Befunde gemäß Abbildung 1 gegenübergestellt und die diagnostische Güte untersucht.

Ein Nutzen lässt sich insgesamt nur dann feststellen, wenn entsprechende Erkenntnisse mindestens zu folgenden Teilfragen vorliegen:

- diagnostische Güte der pränatalen Bestimmung des fetalen Rh(D)-Faktors
- Nutzen und Schaden des Unterlassens der Anti-D-Prophylaxe bei negativem Testergebnis (richtig-negative und falsch-negative Testung des fetalen Rh[D]-Faktors)

In der Subgruppe der Schwangeren mit negativem Testergebnis des fetalen Rh(D)-Faktors könnte ein Nutzen entstehen, insbesondere wenn unerwünschte Wirkungen durch Unterlassen der Anti-D-Prophylaxe vermieden werden können. Umgekehrt könnte ein Schaden entstehen, wenn nach falsch-negativem Test auf eine Anti-D-Prophylaxe verzichtet wird, was dann zu Sensibilisierungen führen könnte.

Es liegt ausreichende Evidenz dafür vor, dass aus einer durch unterlassene, aber indizierte Anti-D-Prophylaxe bewirkten Sensibilisierung gegen das Rh(D)-Antigen ein Schaden resultieren kann (siehe Kapitel 1). Der Nutzen des Unterlassens einer Anti-D-Prophylaxe hinsichtlich vermiedener unerwünschter Wirkungen lässt sich auch ohne Pränataltest an Rh(D)-negativen Schwangeren untersuchen.

Daher sind zusammenfassend mindestens Studienergebnisse zum Unterlassen einer Anti-D-Prophylaxe bei Rh(D)-negativen Schwangeren hinsichtlich vermiedener unerwünschter Wirkungen und zur diagnostischen Güte notwendig, um einen Nutzen des Pränataltests feststellen zu können.

Details zu diesem Vorgehen finden sich in den Abschnitten 4.2 und 4.3.

Untersuchungen zur Subgruppe der Schwangeren mit positivem Testergebnis des Fetus (Effekte aufgrund veränderter Adhärenz) sind ohne Einsatz des Pränataltests nicht möglich und werden gemäß Abschnitt 4.1.7 berücksichtigt.

4.1 Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Untersuchung

4.1.1 Population

Die Zielpopulation der Untersuchung bilden Rh(D)-negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen.

4.1.2 Prüf- und Vergleichsintervention

Die Prüfintervention ist eine nicht invasive molekulargenetische pränatale Bestimmung des Rh(D)-Faktors des Fetus und ein regelhaftes Unterlassen der pränatalen Anti-D-Prophylaxe bei einem Testergebnis, das einen Rh(D)-negativen Fetus anzeigt. Prä- und postpartale Anti-D-Prophylaxe erhalten ausschließlich die Frauen mit einem pränatalen Testergebnis, das einen Rh(D)-positiven Fetus anzeigt.

Die Vergleichsintervention stellt die regelhafte pränatale zulassungsgemäße Anti-D-Prophylaxe bei allen Rh(D)-negativen Schwangeren ohne pränatale Bestimmung des Rh(D)-Faktors dar. Postpartale Anti-D-Prophylaxe erhalten ausschließlich die Frauen, bei denen ein postnataler Test ein Rh(D)-positives Neugeborenes anzeigt.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevanten Endpunkte betrachtet:

- Mortalität
- Auftreten einer hämolytischen Anämie des Fetus beziehungsweise Neugeborenen infolge einer Rh(D)-Inkompatibilität und damit zusammenhängende Komplikationen
- unerwünschte Wirkungen der pränatalen Anti-D-Prophylaxe
- gesundheitsbezogene Lebensqualität

Ist eine Bewertung auf Basis des Endpunktes unerwünschte Wirkungen im Zusammenhang mit einer hämolytischen Anämie eines Fetus beziehungsweise Neugeborenen infolge einer Rh(D)-Inkompatibilität nicht möglich, wird auf folgenden Endpunkt zurückgegriffen:

- Sensibilisierung gegen das Rh(D)-Antigen als ausreichend valides Surrogat für den patientenrelevanten Endpunkt unerwünschte Wirkungen im Zusammenhang mit einer hämolytischen Anämie eines Fetus beziehungsweise Neugeborenen infolge einer Rh(D)-Inkompatibilität

Der Laborwert Sensibilisierung gegen das Rh(D)-Antigen wird als ausreichend valider Surrogatendpunkt betrachtet, da ausreichende Evidenz dafür vorliegt, dass aus einer Sensibilisierung gegen das Rh(D)-Antigen ein Schaden in relevanter Größenordnung resultieren kann, der ohne Sensibilisierung nicht möglich ist (siehe Kapitel 1).

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

4.1.4 Studientypen

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

Für alle unter 4.1.2 genannten Interventionen und alle unter 4.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher primär RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen. Es ist zu erwarten, dass zu dieser Fragestellung hochwertige Kohortenstudien vorliegen. Sollte die auf RCTs basierende Datenlage nicht hinreichend sein, werden auch nicht randomisierte prospektiv geplante vergleichende Interventionsstudien der gesamten diagnostisch-therapeutischen Behandlungskette mit zeitlich paralleler Kontrollgruppe und adäquater Confounderkontrolle zur Nutzenbewertung herangezogen.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.1.6 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 1: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostisch-therapeutischen Behandlungskette)

Einschlusskriterien	
Ea1	Population: Rh(D)-negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen
Ea2	Prüfintervention: nicht invasive molekulargenetische pränatale Bestimmung des Rh(D)-Faktors des Fetus und regelhaftes Unterlassen der pränatalen Anti-D-Prophylaxe bei einem Testergebnis, das einen Rh(D)-negativen Fetus anzeigt (siehe auch Abschnitt 4.1.2)
Ea3	Vergleichsintervention: Gabe der Anti-D-Prophylaxe an alle Rh(D)-negativen Schwangeren (siehe auch Abschnitt 4.1.2)
Ea4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
Ea5	RCTs, gegebenenfalls nicht randomisierte prospektiv geplante Interventionsstudien mit zeitlich paralleler Kontrollgruppe und adäquater Confounderkontrolle (siehe auch Abschnitt 4.1.4)
Ea6	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht oder ein Bericht über die Studie, der den Kriterien des CONSORT- [8] oder TREND-Statements [9] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>Anti-D: gegen das Protein D gerichtetes Immunglobulin; CONSORT: Consolidated Standards of Reporting Trials; Rh(D): Protein D des Rhesus-Blutgruppensystems; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

4.1.7 Studien im Anreicherungsdesign

In die Bewertung fließen auch Studien im Anreicherungsdesign ein, die ausschließlich Effekte bei Schwangeren mit einem bestimmten Testergebnis (Fetus Rh[D]-positiv oder Rh[D]-negativ) untersuchen. Bei Schwangeren mit einem Pränatal-Testergebnis, das einen Rh(D)-positiven Fetus anzeigt, lassen sich die Effekte aufgrund veränderter Adhärenz mit und ohne Kenntnis des Testergebnisses untersuchen. Die Effekte einer veränderten Behandlungsstrategie (mit und ohne Anti-D-Prophylaxe) können an Schwangeren mit Rh[D]-negativ getestetem Fetus erforscht werden. Daher fließen in die Bewertung auch Studien im Anreicherungsdesign ein, die ausschließlich Effekte bei Schwangeren mit einem bestimmten Pränatal-Testergebnis untersuchen.

4.2 Kriterien für den Einschluss von Studien zum Unterlassen einer Anti-D-Prophylaxe in die Untersuchung

4.2.1 Population

Die Zielpopulation der Untersuchung bilden Rh(D)-negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen.

4.2.2 Prüf- und Vergleichsintervention

Die Prüfindervention ist das Unterlassen einer Anti-D-Prophylaxe.

Die Vergleichsintervention ist die zulassungsgemäße Gabe von Anti-D-Prophylaxe.

4.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevanten Endpunkte betrachtet:

- Mortalität
- unerwünschte Wirkungen
- gesundheitsbezogene Lebensqualität

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

4.2.4 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 oder Zusatznutzens einer medizinischen Intervention.

Für alle unter 4.2.2 genannten Interventionen und alle unter 4.2.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher primär RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Es ist denkbar, dass die zu prüfende Intervention einen dramatischen Effekt hinsichtlich unerwünschter Wirkungen aufweist, der sich auch in Studien mit niedrigerem Evidenzniveau nicht allein durch Verzerrung erklären lässt. Wenn die auf RCTs basierende Datenlage nicht reicht, um den patientenrelevanten Nutzen des Unterlassens einer prä- und postpartalen Anti-D-Prophylaxe 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. Auf Basis solcher Studien sind Nutzensaussagen nur möglich, wenn dramatische Effekte vorliegen.

4.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.2.6 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 2: Übersicht über die Kriterien für den Studieneinschluss (Studien zum Unterlassen nicht indizierter Anti-D-Prophylaxe)

Einschlusskriterien	
Eb1	Population: Rh(D)-negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen
Eb2	Prüfintervention: Unterlassen einer Anti-D-Prophylaxe
Eb3	Vergleichsintervention: Gabe einer Anti-D-Prophylaxe
Eb4	patientenrelevante Endpunkte wie in Abschnitt 4.2.3 formuliert
Eb5	RCTs, Kohortenstudien (auch retrospektiv oder mit historischem Vergleich)
Eb6	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des CONSORT- [8], TREND- [9] oder STROBE-Statements [10] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind. Anti-D: gegen das Protein D gerichtetes Immunglobulin; CONSORT: Consolidated Standards of Reporting Trials; Rh(D): Protein D des Rhesus-Blutgruppensystems; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs	

4.3 Kriterien für den Einschluss von Studien zur diagnostischen Güte

4.3.1 Population

Die Zielpopulation der Untersuchung bilden Rh(D)-negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen.

4.3.2 Indextest

Der Indextest ist eine nicht invasive molekulargenetische pränatale Bestimmung des Rh(D)-Faktors des Fetus.

4.3.3 Referenztest

Den Referenztest stellt die postnatale Bestimmung des Rh(D)-Faktors des Kindes dar.

4.3.4 Zielgrößen

Eingeschlossen werden Studien, aus denen personenbezogene Ergebnisse zum prä- und postnatal bestimmten Rh(D)-Faktor zur Berechnung der diagnostischen Güte des Indextestes ableitbar sind.

4.3.5 Studientypen

Um die diagnostische Güte des Indextests zur Bestimmung des Rhesusfaktors des Fetus möglichst unverzerrt bestimmen zu können, fließen prospektive diagnostische Kohortenstudien mit pränataler (Indextest) und postnataler (Referenztest) Rh(D)-Faktor-Bestimmung in die Nutzenbewertung ein. Dabei sind ein konsekutiver Einschluss der Frauen und die Dokumentation fehlender Werte notwendig.

4.3.6 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.3.7 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	
Ec1	Population: Rh(D)- negative Schwangere ohne Sensibilisierung gegen das Rh(D)-Antigen (siehe auch Abschnitt 4.3.1)
Ec2	Indextest: nicht invasive molekulargenetische pränatale Bestimmung des Rh(D)-Faktors des Fetus (siehe auch Abschnitt 4.3.2)
Ec3	Referenztest: postnatale Bestimmung des Rh(D)-Faktors des Kindes (siehe auch Abschnitt 4.3.3)
Ec4	personenbezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt 4.3.4)
Ec5	prospektiv geplante Kohortenstudien wie in Abschnitt 4.3.5 beschrieben
Ec6	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des STARD- [11] oder STROBE-Statements [10] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind. Rh(D): Protein D des Rhesus-Blutgruppensystems; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology	

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

Für das Einschlusskriterium E1 (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 E1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

Ebenfalls eingeschlossen werden Studien, die zu mindestens 80 % das Einschlusskriterium E2 erfüllen (Prüfintervention, bezogen auf die Interventionsgruppe der Studie, beziehungsweise Indextest bei Diagnosestudien) und zu mindestens 80 % das Einschlusskriterium E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie, beziehungsweise Referenztest bei Diagnosestudien).

4.5 Informationsbeschaffung

4.5.1 Primäre Suchquellen

4.5.1.1 Bibliografische Recherche

Die systematische Recherche nach relevanten Studien bzw. Dokumenten 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.

4.5.1.2 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,
- European Medicines Agency. EU Clinical Trials Register.

4.5.2 Weitere Suchquellen

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien beziehungsweise Informationen zu relevanten Studien zu ermitteln, werden weitere Quellen berücksichtigt.

4.5.2.1 Systematische Übersichten

Systematische Übersichten werden hinsichtlich weiterer relevanter Studien bzw. Dokumente gesichtet.

4.5.2.2 Öffentlich zugängliche Dokumente von Zulassungsbehörden

Zusätzlich wird nach öffentlich zugänglichen Dokumenten von Zulassungsbehörden gesucht.

- European Medicines Agency. Website. URL: <http://www.ema.europa.eu>
- Food and Drug Administration. Website. URL: <http://www.fda.gov>

4.5.2.3 Durch den G-BA übermittelte Dokumente

Die vom G-BA mit Auftragserteilung an das IQWiG weitergeleiteten Dokumente werden hinsichtlich weiterer relevanter Studien bzw. Dokumente gesichtet.

4.5.2.4 Anhörung

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.

4.5.2.5 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.

4.5.3 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente 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 1, Tabelle 2 und Tabelle 3) bewertet. Als potenziell relevant erachtete Dokumente 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. Konferenzabstracts werden im Rahmen der Nutzenbewertung nicht berücksichtigt.

Selektion relevanter Studien bzw. Dokumente aus weiteren Suchquellen

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

- öffentlich zugängliche Studienregister,
- öffentlich zugängliche Dokumente von Zulassungsbehörden,
- durch den G-BA übermittelte Dokumente.

Informationen aus den folgenden Suchquellen 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:

- identifizierte systematische Übersichten,

- im Rahmen der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht eingereichte Informationen.

Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden Reviewern aufgelöst. Konferenzabstracts werden im Rahmen der Nutzenbewertung nicht berücksichtigt.

4.6 Informationsbewertung

4.6.1 Bewertung von vergleichenden Interventionsstudien

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.

Datenextraktion

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

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.

4.6.2 Bewertung von Studien zur diagnostischen Güte

Datenextraktion

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

Bewertung des Verzerrungspotenzials der Ergebnisse

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der Primärstudien zur diagnostischen Güte erfolgt auf Basis des QUADAS-2-Instruments [12]. 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.

4.7 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.

4.7.1 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 [13].

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 [14].

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.

4.7.2 Meta-Analysen

4.7.2.1 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 Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Auswahl der Modelle für Meta-Analysen erfolgt gemäß den Kriterien, die in den Allgemeinen Methoden [7] genannt sind. 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 Ratios 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 [15].

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 [16]. 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 4.7.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.7.5).

4.7.2.2 Meta-Analysen für Studien zur diagnostischen Güte

Die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle [17] aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen erfüllt sind, für die Testgütekriterien Sensitivität und Spezifität eine bivariate Meta-Analyse durchgeführt [18]. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [19,20]. 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. In diesem Fall wird für die metaanalytische Zusammenfassung auf den negativen prädiktiven Wert (NPV) zurückgegriffen.

Falls die bivariate Meta-Analyse präzise Schätzungen liefert, so werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Ergebnisse verschiedener Indextests, die aus derselben Studie stammen, werden durch eine Verbindungslinie gekennzeichnet. Des Weiteren werden die aus der bivariaten Meta-Analyse gewonnenen Schätzungen für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95 %-Konfidenzregionen dargestellt [21].

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.

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen untersucht.

4.7.3 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 beurteilen. 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 4 zu entnehmen.

Tabelle 4: 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.						

4.7.4 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 eingestuftter Effekt kann

zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (Zusatz-)Nutzen attestiert wird.

4.7.5 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 4.7.4 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:

- Gestationsalter bei Testdurchführung
- Mehrlingsschwangerschaft

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.

5 Literatur

1. Zimmermann R. Alloimmunerkrankungen und Schwangerschaft. In: Schneider H, Husslein P, Schneider KTM (Ed). Die Geburtshilfe. Berlin: Springer; 2016. S. 615-627.
2. Gemeinsamer Bundesausschuss. Richtlinien des Gemeinsamen Bundesausschusses über die ärztliche Betreuung während der Schwangerschaft und nach der Entbindung („Mutterschafts-Richtlinien“) [online]. 21.04.2016 [Zugriff: 06.10.2016]. URL: https://www.g-ba.de/downloads/62-492-1223/Mu-RL_2016-04-21_2016-07-20.pdf.
3. CSL Behring. Rhophylac 300: Fachinformation [online]. 04.2015 [Zugriff: 05.10.2016]. URL: <http://www.fachinfo.de/>.
4. Octapharma. Rhesonativ: Fachinformation [online]. 09.2015 [Zugriff: 05.10.2016]. URL: <http://www.fachinfo.de/>.
5. Statistisches Bundesamt. Bevölkerung: Deutschland; Lebendgeborene und Gestorbene [online]. [Zugriff: 06.10.2016]. URL: <https://www.destatis.de/DE/ZahlenFakten/Indikatoren/LangeReihen/Bevoelkerung/lrbev04.html;jsessionid=BF9F7E0616671C7A44A20581DCE1A64E.cae4>.
6. National Institute for Health and Care Excellence. New blood test for pregnant women could help thousands avoid unnecessary treatment [online]. 14.07.2016 [Zugriff: 02.09.2016]. URL: <https://www.nice.org.uk/news/article/new-blood-test-for-pregnant-women-could-help-thousands-avoid-unnecessary-treatment>.
7. 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.
8. 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.
9. 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.
10. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
11. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003; 138(1): 40-44.
12. 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. *Annals of Internal Medicine* 2011; 155(8): 529-536.

13. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
14. 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.
15. 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.
16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.
17. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the Bernoulli parameter. *Am Stat* 1996; 50(1): 63-68.
18. 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.
19. 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.
20. 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.
21. Hotelling H. The generalization of student's ratio. *Annals of Mathematical Statistics* 1931; 2(3): 360-378.