

## **Kurzfassung**

Das Bundesministerium für Gesundheit (BMG) hat am 14.09.2016 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Erstellung von Evidenzberichten zur Unterstützung der Entwicklung der S3-Leitlinie „Die vaginale Geburt am Termin“ beauftragt.

## **Fragestellung**

Von den für die Leitlinienerstellung federführenden Fachgesellschaften, der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG) und der Deutschen Gesellschaft für Hebammenwissenschaft e. V. (DGHWi), wurden zusammen mit der AWMF 8 Participants-Intervention-Comparison-Outcomes(PICO)-Fragen formuliert, zu denen das IQWiG Evidenzberichte im Format von Rapid Reports erstellen soll.

Die Evidenz für folgende PICO-Fragen wurde in diesem Evidenzbericht zusammengefasst:

- Wie wirkt sich eine intrapartale Ultraschalldiagnostik ab Aufnahme in den Kreißsaal auf verschiedene maternale und neonatale Endpunkte im Vergleich zum Verzicht auf diese Diagnostik aus?
- Welche Testgüte zeigt die Ultraschalldiagnostik in Bezug auf das Erkennen von Plazentapathologien, die Bestimmung des kindlichen Gewichts und / oder Wachstumsauffälligkeiten?

Die intrapartale Ultraschalldiagnostik ist für diesen Bericht gemäß Auftrag definiert als eine Ultraschalluntersuchung bei Aufnahme und / oder während der Geburt (z. B. eine Fetometrie, eine Lagekontrolle des Fetus oder der Einstellung des Kopfes, eine Beurteilung des Plazentasitzes und der -reife, eine Doppleruntersuchung zur Bestimmung der fetalen Versorgung oder ein transvaginaler Ultraschall zur Diagnostik des Muttermundbefunds).

## **Methoden**

Die Zielpopulation dieser Evidenzbewertung bildeten Schwangere um den errechneten Geburtstermin ab vollendeter 37. bis zur vollendeten 42. Schwangerschaftswoche (37 + 0 bis 42 + 0) ab Aufnahme in den Kreißsaal. Ein intrapartaler Ultraschall soll über die Früherkennung eines auffälligen Befunds die Möglichkeit bieten, therapeutische Konsequenzen einzuleiten und das Management während der Geburt anzupassen.

Die von der Leitliniengruppe festgelegten Endpunkte waren gemäß der Vorgehensweise nach dem GRADE-Ansatz in die Kategorien „kritisch für die Entscheidung“ (maternale Endpunkte: Mortalität, Geburtsmodus [Spontangeburt, vaginal-operative Geburt oder Kaiserschnitt], höhergradige Dammriss [III°, IV°], hoher mütterlicher Blutverlust  $\geq 1000$  ml, Aufnahme einer Stillbeziehung bei initialem Stillwunsch, teilweise oder vollständige Plazentaretention, Geburtserleben, subjektive Gesundheitseinschätzung; neonatale Endpunkte: Mortalität, Apgar-Score nach 5 Minuten  $< 6$ , schwere metabolische Acidose [pH  $< 7,0$ ],

Basendefizit  $\geq -16$  mmol/l, Verlegung auf die Neugeborenen-Intensivstation) und „klinisch wichtig, aber für die Empfehlung nicht entscheidend“ (maternale Endpunkte: Episiotomie, Einsatz von Wehenmitteln, Amniotomie, Kristellerhilfe, Infektionen; neonatale Endpunkte: leichte, mittelgradige oder fortgeschrittene Acidose, Ikterus, Anämie) zugeordnet.

Im 1. Schritt wurden randomisierte kontrollierte Studien im Strategiedesign eingeschlossen, in denen Managementstrategien mit und ohne Verwendung des intrapartalen Ultraschalls oder Gruppen mit beziehungsweise ohne Offenlegung der diagnostischen Testergebnisse verglichen wurden. In einem 2. Schritt wurden randomisierte kontrollierte Studien im Anreicherungsdesign zur Untersuchung von Therapieeffekten möglicher Managementoptionen bei auffälligem Ultraschallbefund eingeschlossen. Die Suche erfolgte fokussiert auf Wachstumsauffälligkeiten und Plazentapathologien. Zudem sollten systematische Übersichten zur Testgüte des Ultraschalls zur Genauigkeit der Schätzung des kindlichen Gewichts und /oder zum Erkennen von Wachstumsauffälligkeiten und Plazentapathologien eingeschlossen werden. Für alle Studien wurde eine Zeitbeschränkung auf die Jahre ab 1990 festgelegt.

Eine systematische Literaturrecherche nach Primärstudien wurde in den Datenbanken MEDLINE, Embase, Cochrane Central Register of Controlled Trials und Cinahl durchgeführt. Parallel erfolgte eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE, Embase, Cinahl, Cochrane Database of Systematic Reviews und HTA Database.

Darüber hinaus wurden folgende Informationsquellen und Suchtechniken berücksichtigt: Studienregister, die Sichtung von Referenzlisten, Anfragen an Studiengruppen und Fachgesellschaften sowie Autorenanfragen.

Des Weiteren wurde eine fokussierte Informationsbeschaffung bezüglich systematischer Übersichten zur Testgüte in den Datenbanken MEDLINE, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews und HTA Database durchgeführt.

Die Selektion relevanter Studien erfolgte von 2 Reviewerinnen unabhängig voneinander, Diskrepanzen wurden durch Diskussion aufgelöst. Die Datenextraktion erfolgte in standardisierte Tabellen. Das Verzerrungspotenzial wurde für die randomisierten kontrollierten Studien auf Studien- und Endpunktebene bewertet und jeweils als gering, mittel oder hoch eingestuft. Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren und keine bedeutsame Heterogenität beobachtet wurde, wurden die Einzelergebnisse für die kritischen Endpunkte mithilfe von Metaanalysen quantitativ zusammengefasst. Zusätzlich erfolgten Subgruppenanalysen bezogen auf Gebärende entsprechend der Parität. Die dabei berechneten relativen Risiken bildeten unter Berücksichtigung der Populationsrisiken die Grundlage für die Berechnung absoluter Effekte.

Für jeden kritischen Endpunkt wurde anschließend eine Bewertung der Qualität der Evidenz nach der GRADE-Methodik in 4 Abstufungen (hoch, moderat, niedrig oder sehr niedrig) bezüglich des Vertrauens in die Effektschätzungen vorgenommen. Anschließend erfolgte eine Gesamtbewertung der Qualität der Evidenz.

Die Bewertung der Testgüte sollte auf systematischen Übersichten hoher methodischer Qualität basieren. Die Charakteristika der dargestellten systematischen Übersichten, Informationen zum Verzerrungspotenzial der eingeschlossenen Studien und Ergebnisse zur Sensitivität und Spezifität wurden ergänzend extrahiert.

### ***Ergebnisse der Informationsbeschaffung***

Die Informationsbeschaffung identifizierte 6 randomisierte kontrollierte Studien (11 Dokumente) zu Managementstrategien mit und ohne intrapartalen Ultraschall und 3 randomisierte kontrollierte Studien (4 Dokumente) zu verschiedenen Managementoptionen bei auffälligem Befund bezogen auf die fetale Gewichtsschätzung beziehungsweise Wachstumsauffälligkeiten als relevant für die Fragestellung der vorliegenden Bewertung.

Es konnte keine systematische Übersicht zur Testgüte bezüglich der Gewichtsschätzung und/oder des Erkennens von Wachstumsauffälligkeiten beziehungsweise von Plazentopathologien eines intrapartal oder ab der vollendeten 37. Schwangerschaftswoche durchgeführten Ultraschalls identifiziert werden. Deshalb wurden 4 systematische Übersichten mit breiteren Einschlusskriterien hinsichtlich ihrer methodischen Qualität bewertet und deren Ergebnisse ergänzend dargestellt.

Es wurden 9 laufende Studien (5 zu Managementstrategien mit und ohne intrapartalen Ultraschall, 4 zu verschiedenen Managementoptionen bei auffälligem Befund) identifiziert. Des Weiteren wurde 1 abgeschlossene Studie zu Managementstrategien ohne berichtete Ergebnisse identifiziert. Die letzte Suche fand am 27.07.2017 statt.

### ***Ergebnisse der in die Bewertung eingeschlossenen Studien beziehungsweise ergänzend dargestellten systematischen Übersichten***

Zu Managementstrategien mit und ohne intrapartalen Ultraschall wurden insgesamt 6 randomisierte kontrollierte Studien zu 5 verschiedenen Einsatzmöglichkeiten identifiziert.

1 Studie untersuchte unter Einschluss von 965 Frauen mit bestimmten Risikofaktoren und Schwangerschaftskomplikationen ab der vollendeten 37. Schwangerschaftswoche die Wirksamkeit einer seriellen ultrasonografischen fetalen Gewichtsschätzung im 3. Trimenon zur präpartalen Diagnose von Small-for-gestational-age-Feten und den daraus folgenden therapeutischen Konsequenzen. Bei einer insgesamt sehr niedrigen Qualität der Evidenz bleibt unklar, ob sich Managementstrategien, die auf einer Gewichtsschätzung mittels Ultraschalluntersuchung basieren positiv oder negativ auf die berichteten kritischen Endpunkte (Geburtsmodus, neonatale Mortalität, Verlegung auf die Neugeborenen-Intensivstation) auswirken.

1 Studie untersuchte unter Einschluss von 58 Erstgebärenden mit Feten in occipito-posteriorer Einstellung und Geburtsstillstand in der Austreibungsperiode die Wirksamkeit einer transabdominalen Ultraschalluntersuchung vor dem Versuch der manuellen Rotation. Diese führte zu mehr Spontangeburt (RR: 2,5; 98,3 %-KI: [1,15; 5,45]) und weniger vaginal-operativen Geburten (RR: 0,44; 98,3 %-KI: [0,20; 0,99]) bei sehr unpräziser Schätzung der Auswirkungen auf die Anzahl der Kaiserschnitte. Bei weiteren kritischen Endpunkten (hoher mütterlicher Blutverlust > 1000 ml, Apgar-Score nach 5 Minuten < 6) traten keine Ereignisse auf. Die Qualität der Evidenz wurde zusammenfassend für Erstgebärende als niedrig und für Mehrgebärende als sehr niedrig bewertet.

1 Studie untersuchte unter Einschluss von 1903 Frauen ab einer Muttermundseröffnung  $\geq 8$  cm die Wirksamkeit einer transabdominalen Ultraschalluntersuchung zur Bestimmung des Höhenstands des fetalen Kopfes. Diese führte bei insgesamt moderater Qualität der Evidenz zu weniger Spontangeburt (RR: 0,91; 98,3 %-KI: [0,85; 0,98]) und mehr Kaiserschnittgeburten (RR: 1,60; 98,3 %-KI: [1,04; 2,47]). Ob sich der Ultraschall und daraus folgende Managementstrategien auf die Anzahl vaginal-operativer Geburten und weitere kritische Endpunkte (höhergradige Dammriss, Verlegung auf die Neugeborenen-Intensivstation) positiv oder negativ auswirkt, bleibt unklar.

2 Studien untersuchten unter Einschluss von insgesamt 564 Gebärenden, bei denen eine vaginal-operative Geburt notwendig war, die Wirksamkeit einer zusätzlichen transabdominalen Ultraschalluntersuchung zur Bestimmung des Höhenstands des fetalen Kopfes. Ob sich der Ultraschall und daraus folgende Managementstrategien auf die berichteten kritischen Endpunkte (Geburtsmodus, höhergradige Dammriss, neonatale Mortalität, Verlegung auf die Neugeborenen-Intensivstation) positiv oder negativ auswirkt, bleibt unklar. Zusammenfassend wurde die Qualität der Evidenz für Erstgebärende als niedrig und für Mehrgebärende als sehr niedrig bewertet.

1 Studie untersuchte unter Einschluss von 90 Mehrgebärenden die Wirksamkeit einer transperinealen Ultraschalluntersuchung im Vergleich zu einer digitalen vaginalen Untersuchung zur Erhebung des vaginalen Befunds während der Geburt. Bei einer insgesamt sehr niedrigen Qualität der Evidenz bleibt unklar, ob der Ultraschall und daraus folgende Managementstrategien die Anzahl von Kaiserschnittgeburten erhöht oder senkt. Weitere kritische Endpunkte wurden nicht berichtet.

Zum Vergleich verschiedener Managementoptionen bei auffälligem Befund wurden insgesamt 3 randomisierte kontrollierte Studien bei 2 auffälligen Befunden (Verdacht auf intrauterine Wachstumsretardierung [IUGR] und Verdacht auf Makrosomie / Large for gestational age [LGA]) identifiziert, die eine zeitnahe Einleitung der Geburt mit einem initial abwartenden Management bis zum Beginn der spontanen Wehentätigkeit verglichen. Es konnte keine Studie zum Vergleich verschiedener Managementoptionen bei Verdacht auf Plazentopathologien eingeschlossen werden.

1 Studie untersuchte unter Einschluss von 33 Schwangeren mit Verdacht auf IUGR die Wirksamkeit beider Strategien. Bei einer insgesamt sehr niedrigen Qualität der Evidenz bleibt unklar, ob sich die Änderung des Managements nach Diagnose eines auffälligen Ultraschallbefunds positiv oder negativ auf die berichteten kritischen Endpunkte (Geburtsmodus, Apgar-Score nach 5 Minuten  $< 6$ , schwere metabolische Acidose, Verlegung auf die Neugeborenen-Intensivstation) auswirkt.

2 Studien untersuchten unter Einschluss von 1106 Schwangeren mit Verdacht auf Makrosomie / LGA die Wirksamkeit beider Strategien. Bei einer insgesamt sehr niedrigen Qualität der Evidenz bleibt unklar, ob sich die Änderung des Managements nach Diagnose eines auffälligen Ultraschallbefunds positiv oder negativ auf die berichteten kritischen Endpunkte (Geburtsmodus, höhergradige Dammriss, hoher mütterlicher Blutverlust  $> 1000$  ml, teilweise / vollständige Plazentaretention, neonatale Mortalität, schwere metabolische Acidose, Verlegung auf die Neugeborenen-Intensivstation) auswirkt.

Die Testgüte des Ultraschalls zur Genauigkeit der Schätzung des kindlichen Gewichts und / oder zum Erkennen von Wachstumsauffälligkeiten und Plazentopathologien bleibt unklar, da keine geeigneten systematischen Übersichten eingeschlossen werden konnten. Es wurden die Ergebnisse aus jeweils 2 systematischen Übersichten mit unterschiedlichen Zielsetzungen ergänzend dargestellt, welche Studien zur Testgüte einschlossen, die Frauen bereits vor Aufnahme in den Kreißsaal rekrutierten und deren Übertragbarkeit auf die Fragestellung dieses Rapid Reports nur eingeschränkt möglich ist. Aufgrund der unzureichenden Datenlage zur Testgüte der Ultraschalldiagnostik zum Erkennen von Wachstumsauffälligkeiten und Plazentopathologien sowie aufgrund der unklaren Vor- und Nachteile von Managementanpassungen bei diesbezüglich auffälligem Befund kann keine Aussage zu Vor- und Nachteilen des Ultraschalleinsatzes bezüglich kritischer oder wichtiger maternaler und neonataler Endpunkte getroffen werden.

### ***Fazit***

Auf der Basis von 6 randomisierten kontrollierten Studien zu Managementstrategien mit vs. ohne intrapartalem Ultraschall konnte für 2 von 5 untersuchten Indikationen eine Auswirkung der Anpassung des Managements nach einem intrapartalen Ultraschall auf kritische Endpunkte nachgewiesen werden. Ein intrapartaler Ultraschall zur Beurteilung des Höhenstands und der Lage des Rückens bei fetaler occipito-posteriorer Einstellung vor dem Versuch einer manuellen Rotation führte zu einem statistisch signifikanten Anstieg der Häufigkeit von Spontangeburt und einer Abnahme der Anzahl vaginal-operativer Geburten bei einer zusammenfassenden niedrigen Qualität der Evidenz für Erstgebärende und sehr niedrigen Qualität für Mehrgebärende. Ein intrapartaler Ultraschall zur Bestimmung des Höhenstands des fetalen Kopfes während der Geburt führte zu einem statistisch signifikanten Anstieg der Häufigkeit von Kaiserschnitten bei einer Abnahme der Anzahl der Spontangeburt und einer zusammenfassenden moderaten Qualität der Evidenz. Die Bewertungen der zusammenfassenden Qualität der Evidenz des Evidenzkörpers zu weiteren Einsatzmöglichkeiten reichten von sehr niedrig bis niedrig. Die Abstufungen der Evidenz

basierten vorrangig auf schwerwiegenden Studienlimitationen wie einer nicht verblindeten Endpunkterhebung, einem intransparenten Teilnehmerinnenfluss, Unklarheiten bezüglich der Festlegung der Fallzahl sowie einer ergebnisunabhängigen Berichterstattung, einer unzureichenden Präzision der Effektschätzungen und einer Indirektheit für Mehrgebärende.

Es konnten keine prospektiv geplanten vergleichenden Studien identifiziert werden, die eine routinemäßige ultrasonografische Untersuchung bei der Aufnahme in den Kreißsaal (z. B. aufgrund eines vorzeitigen Blasensprungs, bei fraglichem Geburtsbeginn oder zur Einleitung) mit dem Verzicht auf diese Intervention verglichen und den Einschlusskriterien entsprachen.

Bei einem auffälligen Befund (Verdacht auf IUGR bzw. Makrosomie / LGA) bleibt der Einfluss von Managementoptionen auf die kritischen Endpunkte bei Vorliegen von Evidenz aus 3 randomisierten kontrollierten Studien unklar. Zusammenfassend wurde die Qualität der Evidenz der Studien zu Managementoptionen bei Verdacht auf IUGR bzw. Makrosomie / LGA als sehr niedrig bewertet. Die Abstufungen der Evidenz basierten vorrangig auf schwerwiegenden Studienlimitationen wie einer nicht verblindeten Endpunkterhebung und einer unklaren ergebnisunabhängigen Berichterstattung, einer unzureichenden Präzision der Effektschätzungen und einem Publikationsbias bei Verdacht auf Makrosomie / LGA.

Die Testgüte des Ultraschalls zur Schätzung des kindlichen Gewichts und/oder zum Erkennen von Wachstumsauffälligkeiten beziehungsweise Plazentapathologien konnte nicht bewertet werden, da keine systematischen Übersichten hierzu eingeschlossen werden konnten. Die Ergebnisse der ergänzend dargestellten systematischen Übersichten zur Bewertung der Testgüte zur Gewichtsschätzung oder zum Erkennen von Plazentapathologien sind nur begrenzt auf die Zielpopulation übertragbar. Der Einfluss von Managementänderungen bei auffälligem Ultraschallbefund bezogen auf Wachstumsauffälligkeiten auf kritische Endpunkte ist unklar. Daher bliebe selbst bei hinreichender Testgüte der Einfluss eines Ultraschalls zur Managementanpassung bei Wachstumsauffälligkeiten ab der vollendeten 37. SSW unklar.

Ingesamt ist zu beachten, dass mehrere vordefinierte Endpunkte nicht untersucht werden konnten, da nicht für alle Endpunkte Evidenz aus den eingeschlossenen Studien vorlag.