

## **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 bei der 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 Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (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-Frage wurde in diesem Evidenzbericht zusammengefasst:

Wie wirkt sich die Erfahrung / Expertise von ärztlichen Geburtshelferinnen oder Geburtshelfern und Hebammen beziehungsweise Entbindungspflegern bei Schwangeren am Termin unter der Geburt ab der aktiven Eröffnungsperiode (EP) auf verschiedene maternale und neonatale Endpunkte im Vergleich zu dem jeweiligen Fachpersonal mit geringerer Erfahrung / Expertise aus?

## **Methoden**

Die Zielpopulation dieser Evidenzbewertung bildeten Schwangere um den errechneten Geburtstermin ab der vollendeten 37. bis zur vollendeten 42. Schwangerschaftswoche (37 + 0 bis 42 + 0) von der aktiven EP bis zum Ende der Nachgeburtsperiode. Es wurden prospektiv geplante vergleichende Studien mit einem Vergleich einer höheren und einer geringeren Ausprägung von Erfahrung / Expertise von ärztlichen Geburtshelferinnen oder Geburtshelfern, Hebammen oder Entbindungspflegern in Krankenhäusern als Prüf- und Vergleichsintervention eingeschlossen. Für alle Studien wurde eine Zeitbeschränkung auf die Jahre ab 1990 festgelegt.

Die von der Leitliniengruppe festgelegten Endpunkte waren gemäß der Vorgehensweise nach dem GRADE-Ansatz (GRADE: Grading of Recommendations, Assessment, Development, and Evaluation) den Kategorien „kritisch für die Entscheidung“ (im Folgenden als „kritisch“ bezeichnet) und „klinisch wichtig, aber für die Empfehlung nicht entscheidend“ (im Folgenden als „wichtig“ bezeichnet) zugeordnet. Der Kategorie der kritischen Endpunkte wurden die 8 maternalen Endpunkte Mortalität, Geburtsmodus mit den 3 Ausprägungen 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 und subjektive Gesundheitseinschätzung zugeordnet. Von den neonatalen Endpunkten zählen die 5 Endpunkte Mortalität, Apgar-Score nach 5 Minuten  $< 6$ , schwere metabolische Acidose ( $\text{pH} < 7,0$ ), Basendefizit  $\geq -16$  mmol/l und Verlegung auf die Neugeborenen-Intensivstation zu der

Kategorie der kritischen Endpunkte. Die 5 maternalen Endpunkte Episiotomie, Einsatz von Wehenmitteln, Amniotomie, Kristellerhilfe und Infektionen wurden der Kategorie wichtig zugeordnet. Von den neonatalen Endpunkten wurden die 3 Endpunkte leichte, mittelgradige oder fortgeschrittene Acidose, Ikterus und Anämie dieser Kategorie zugeordnet.

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.

Die Selektion relevanter Studien erfolgte von 2 Reviewerinnen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen den beiden Reviewerinnen aufgelöst. Die Datenextraktion erfolgte in standardisierte Tabellen. Das Verzerrungspotenzial wurde für die vergleichenden Interventions- und Beobachtungsstudien auf Studien- und Endpunktebene bewertet und konnte als gering, mittel oder hoch eingestuft werden. Es war geplant, die Einzelergebnisse vergleichbarer Studien für die kritischen Endpunkte mithilfe von Metaanalysen quantitativ zusammenzufassen.

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.

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

Die Informationsbeschaffung identifizierte 2 randomisierte kontrollierte Studien (2 Dokumente) und 6 Beobachtungsstudien (6 Dokumente) zum Vergleich größerer Expertise versus geringere Expertise als relevant für die Fragestellung der vorliegenden Bewertung. Es wurde 1 laufende Studie identifiziert. Die letzte Suche fand am 1.10.2018 statt.

### ***Ergebnisse der in die Bewertung eingeschlossenen Studien***

#### *Ergebnisse zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen*

Zur Bewertung der Erfahrung / Expertise wurden 2 randomisierte kontrollierte Studien eingeschlossen, welche Weiterbildungsmaßnahmen des geburtshilflichen Fachpersonals vergleichen. Die 2 Studien wurden in Kanada in den Jahren 1988 bis 1989 und 2008 bis 2011 durchgeführt.

In 1 Studie wurden Daten von insgesamt 52 265 Erst- und Mehrgebärenden und 53 113 Neugeborenen nach Durchführung der Weiterbildungsmaßnahmen ausgewertet. Das Ziel der Studie war die Evaluation eines Programms zur Schulung des geburtshilflichen

Personals. In 16 Krankenhäusern beinhaltete die Prüfintervention eine 6-monatige Trainingsphase mit einer Schulung zur evidenzbasierten klinischen Praxis, der Wahl eines Meinungsführers und eines Auditausschusses durch das geburtshilfliche Personal und die Durchführung von Workshops zu Audit- und Feedback-Maßnahmen und zur Verbesserung der Indikationsstellung eines intrapartalen Kaiserschnitts sowie des intrapartalen Managements. Abschließend erfolgten 4 durch den Auditausschuss organisierte Auditzyklen über jeweils 3 Monate mit externer Unterstützung. Während des 1-jährigen Zeitraums nach Durchführung der Weiterbildungsmaßnahme wurde das Gesundheitsfachpersonal ermutigt, die klinischen Audits auch ohne Supervision weiterzuführen. In weiteren 16 Krankenhäusern wurden als Vergleichsintervention keine Weiterbildungsmaßnahmen durchgeführt.

In der 2. Studie wurden Daten von insgesamt 914 Mehrgebärenden mit einem vorherigen Kaiserschnitt und einer angestrebten vaginalen Geburt aus 16 Krankenhäusern ausgewertet. Das Ziel der 3-armigen Studie war die Evaluation von 2 Strategien zur Weiterbildung des ärztlichen Fachpersonals zur verbesserten Umsetzung einer Leitlinie zur Geburtsplanung bei Schwangeren mit Zustand nach Kaiserschnitt. In 4 Krankenhäusern beinhaltete die Prüfintervention eine Schulung durch spezifisch geschulte und in den Einrichtungen anerkannte ärztliche Geburtshelfer („Schulung von Meinungsführer“) und in weiteren 4 Krankenhäusern beinhaltete sie Audit-und-Feedback-Maßnahmen. In 8 Krankenhäusern wurde als Vergleichsintervention dem ärztlichen Fachpersonal eine Kopie der umzusetzenden Leitlinie mit Hinweis auf die relevanten Abschnitte zugesendet.

Es konnten Ergebnisse für insgesamt 5 kritische (maternale und neonatale Mortalität, Geburtsmodus, höhergradiger Dammriss [III°, IV°], schwere metabolische Acidose [pH < 7,0]) und 5 wichtige Endpunkte (Episiotomie, Wehenmittel, Amniotomie, Infektionen, fortgeschrittene metabolische Acidose [pH 7,09-7,0]) ermittelt werden. Die Qualität der Evidenz wurde als niedrig bewertet. Es lagen keine Ergebnisse aus randomisierten kontrollierten Studien zu den kritischen Endpunkten hoher mütterlicher Blutverlust  $\geq 1000$  ml, Aufnahme einer Stillbeziehung bei initialem Stillwunsch, teilweise / vollständige Plazentaretention, Geburtserleben, subjektive Gesundheitseinschätzung, Apgar-Score nach 5 Minuten < 6, Basendefizit  $\geq -16$  mmol/l oder Verlegung auf die Neugeborenen-Intensivstation und weiteren klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte vor.

Bei einer insgesamt niedrigen Qualität der Evidenz führte die Durchführung von Weiterbildungsmaßnahmen zu mehr vaginalen Geburten (OR: 1,07; 97,5 %-KI: [1,00; 1,14]). Eine Abnahme der ungeplanten Kaiserschnitte (OR: 0,89; 97,5 %-KI: [0,78; 1,01]) konnte nicht nachgewiesen werden. Die Veränderungen im Geburtsmodus ging einher mit einer Abnahme der neonatalen Mortalität (OR: 0,34; 95 %-KI: [0,18; 0,65]), ohne dass Auswirkungen auf das Auftreten einer schweren metabolischen Acidose (pH < 7,0) (OR: 1,13; 95 %-KI: [0,79; 1,62]) nachgewiesen werden konnten. Es konnte kein Unterschied für die weiteren berichteten kritischen maternalen Endpunkte maternale Mortalität (OR: 2,08; 95 %-KI: [0,05; 81,6]) und

das Auftreten höhergradiger Dammmrisse (III°, IV°) (OR: 1,07; 95 %-KI: [0,90; 1,26]) nachgewiesen werden.

Für die wichtigen neonatalen Endpunkte waren Weiterbildungsmaßnahmen mit einem zunehmenden Auftreten einer fortgeschrittenen metabolischen Acidose (pH 7,09-7,0) (OR: 1,49, 95 %-KI: [1,02; 2,17]) assoziiert. Bei den wichtigen maternalen Endpunkten wurden Auswirkungen auf die Gabe von Wehenmitteln nachgewiesen: Während weniger Frauen Wehenmittel zur Einleitung der Geburt erhielten (OR: 0,82, 95 %-KI: [0,76; 0,87]), erhielten mehr Gebärende Wehenmittel unter der Geburt (OR: 1,16, 95 %-KI: [1,09; 1,23]). Ein Einfluss auf weitere berichtete wichtige Endpunkte (Episiotomie, Amniotomie und Infektionen) konnte nicht nachgewiesen werden.

Die Herabstufung der Qualität der Evidenz erfolgte vorrangig aufgrund schwerwiegender Studienlimitationen wie einer unklaren oder fehlenden Verblindung der Gebärenden, der behandelnden und endpunkterhebenden Personen und einer Berechnung von Effektschätzern ohne Berücksichtigung der Clusterstruktur, Indirektheit da die berechneten Effekte auf allen Gebärenden, einschließlich der mit geplanten Kaiserschnitt basieren, sowie einer schwerwiegenden fehlenden Genauigkeit.

#### *Ergebnisse zur Erfahrung / Expertise durch systemimmanente Faktoren*

Zur Bewertung der Erfahrung / Expertise durch systemimmanente Faktoren wurden 6 Beobachtungsstudien eingeschlossen, von denen 3 Studien auf Daten der französischen Perinataldatenerhebung und 3 Studien auf nationalen Geburtsregistern basieren. Da keine randomisierten Studien identifiziert werden konnten, stammen alle berichteten Ergebnisse aus diesen prospektiv geplanten vergleichenden Kohortenstudien, in denen nicht verschiedene Interventionen, sondern Expositions- mit Referenzgruppen verglichen werden.

Insgesamt liegen für 1 891 204 Gebärende Ergebnisse vor, die einen Vergleich einer größeren Erfahrung / Expertise und geringeren Expertise auf der Grundlage der Geburtenzahl pro Jahr oder der Versorgungsstufe der Geburtskliniken hinsichtlich ihres Einflusses auf die Endpunkte von Mutter und Kind erlauben. Da metaanalytische Zusammenfassungen der Ergebnisse der Einzelstudien nicht möglich waren, werden die Ergebnisse aus den primär interessierenden Vergleichen zwischen den größten und kleinsten Ausprägungen von Erfahrung / Expertise der Einzelstudien interpretiert, soweit diese auf Deutschland übertragbar sind.

Für den Vergleich einer größeren Erfahrung / Expertise aufgrund einer höheren Geburtenzahl in den Geburtskliniken konnten Ergebnisse für insgesamt 3 kritische Endpunkte (Geburtsmodus, neonatale Mortalität, schwere metabolische Acidose [pH < 7,0]) ermittelt werden. Die Qualität der Evidenz wurde für alle berichteten Endpunkte als sehr niedrig bewertet. Ergebnisse aus prospektiv geplanten Kohortenstudien zur maternalen Mortalität, höhergradigen Dammrissen (III°, IV°), hohem mütterlichen Blutverlust  $\geq 1000$  ml, zur Aufnahme einer Stillbeziehung bei initialem Stillwunsch, zur teilweisen / vollständigen Plazentaretention, zum Geburtserleben, zur subjektiven Gesundheitseinschätzung und zum Basendefizit  $\geq -16$  mmol/l

wurden nicht berichtet. Zu den kritischen Endpunkten Apgar-Score nach 5 Minuten < 6 und Verlegung auf die Neugeborenen-Intensivstation lagen ausschließlich ergänzende Informationen vor. Für die klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte liegen ausschließlich Ergebnisse zu Episiotomie und Wehenmittel vor.

Eine größere Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, führte bei insgesamt sehr niedriger Qualität der Evidenz bei Erstgebärenden zu mehr vaginalen Geburten (OR: 1,22; 97,5 %-KI: [1,14; 1,30]) und weniger ungeplanten Kaiserschnitten (OR: 0,82; 98,3 %-KI: [0,77; 0,88]) und bei Erst- und Mehrgebärenden zu mehr vaginal-operativen Geburten (OR: 1,47; 98,3 %-KI: [1,03; 2,09]) ohne Informationen zur Auswirkung auf die Ausprägungen Spontangeburt oder Kaiserschnitt. Die beim Vergleich zwischen großen Geburtskliniken mit > 3000 Geburten pro Jahr und mittelgroßen Kliniken mit 1100 bis 3000 Geburten pro Jahr beobachteten Unterschiede bei vaginalen Geburten und Kaiserschnitten bei Erstgebärenden konnten beim Vergleich zwischen mittelgroßen und kleineren Kliniken mit < 1100 Geburten pro Jahr nicht bestätigt werden (OR: 1,02; 97,5 %-KI: [0,91; 1,15] für vaginale Geburt und OR: 0,98; 98,3 %-KI: [0,86; 1,12] für Kaiserschnitt). Bei den neonatalen Endpunkten war eine größere Geburtenzahl pro Jahr in der Geburtsklinik mit einem häufigeren Auftreten einer schweren metabolischen Acidose (OR: 1,43; 95 %-KI: [1,14; 1,79]) assoziiert. Diese führte aber zu keinem Unterschied hinsichtlich der intrapartalen (OR: 0,84; 95 %-KI: [0,66; 1,05]) und neonatalen Mortalität (OR: 0,47; 95 %-KI: [0,22; 1,03] an den Tagen 1–7 und OR: 1,19; 95 %-KI: [0,90; 1,59] an den Tagen 1–27). Für die wichtigen Endpunkte war eine Entbindung in einem Krankenhaus mit größeren Geburtenzahlen mit einem geringeren Einsatz von Wehenmitteln (OR: 0,67; 95 %-KI: [0,56; 0,83]) bei Gebärenden mit einer PDA assoziiert, ein Einfluss auf die Anzahl der Episiotomien bei Gebärenden mit einer Spontangeburt konnte nicht nachgewiesen werden (OR: 0,80; 95 %-KI: [0,57; 1,12]).

Die Herabstufung der Qualität der Evidenz basiert vorrangig auf dem Studiendesign der eingeschlossenen nicht randomisierten Studien und weiteren Abstufungen aufgrund von schwerwiegenden Studienlimitationen, schwerwiegender Indirektheit für Mehrgebärende und schwerwiegender fehlender Genauigkeit. Die schwerwiegenden Studienlimitationen umfassen eine unklare Konfounderadjustierung bei fehlenden Informationen zu Baseline-Charakteristika in den Expositionsgruppen für die neonatalen Endpunkte, eine fehlende Verblindung der teilnehmenden und behandelnden Personen, eine unklare oder fehlende ergebnisunabhängige Berichterstattung, eine möglicherweise unterschiedliche Erfassung von Risikofaktoren in den Datenbanken und den Ausschluss von Neugeborenen aus der Analyse aufgrund fehlender Daten und die daraus resultierende mögliche Beeinflussung des Ergebnisses für seltene Ereignisse.

Für den Vergleich einer größeren Erfahrung / Expertise aufgrund einer höheren Versorgungsstufe der Geburtskliniken konnten Ergebnisse für insgesamt 1 kritischen Endpunkt (Geburtsmodus) zur Ausprägung vaginal-operativer Geburt ermittelt werden. Ergebnisse aus prospektiv geplanten Kohortenstudien zu weiteren kritischen Endpunkten wurden nicht berichtet. Für die klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte liegen ausschließlich Ergebnisse zu Episiotomie und Wehenmittel vor.

Für den Einfluss einer größeren Erfahrung / Expertise, beschrieben durch die Versorgungsstufe der Geburtsklinik, konnte bei insgesamt sehr niedriger Qualität der Evidenz für den Vergleich der Versorgungsstufen Level 3 vs. Level 1 kein Unterschied auf die Anzahl von vaginal-operativen Geburten nachgewiesen werden (OR: 1,30; 98,3 %-KI: [0,95; 1,77]). Für eine abschließende Bewertung des Einflusses der Versorgungsstufe auf den Endpunkt Geburtsmodus fehlen Informationen zu den beiden Ausprägungen Spontangeburt und Kaiserschnitt. Für die wichtigen Endpunkte war eine Entbindung in einem Krankenhaus einer höheren Versorgungsstufe mit einem geringeren Einsatz von Wehenmitteln bei Gebärenden mit einer PDA (OR: 0,63; 95 %-KI: [0,51; 0,77]) assoziiert, ein Einfluss auf die Anzahl der Episiotomien konnte nicht nachgewiesen werden (RR: 1,11; 95 %-KI: [0,91; 1,36]).

Die Herabstufung der Qualität der Evidenz basiert vorrangig auf dem Studiendesign der eingeschlossenen nicht randomisierten Studien und weiteren Abstufungen aufgrund von schwerwiegenden Studienlimitationen, schwerwiegender Indirektheit und schwerwiegender fehlender Genauigkeit. Die schwerwiegenden Studienlimitationen umfassen eine unklare Konfounderadjustierung bei fehlenden Informationen zu Baseline-Charakteristika in den Expositionsgruppen für die neonatalen Endpunkte, eine fehlende Verblindung der teilnehmenden und behandelnden Personen, eine unklare oder fehlende ergebnisunabhängige Berichterstattung, eine möglicherweise unterschiedliche Erfassung von Risikofaktoren in den Datenbanken und den Ausschluss von Neugeborenen aus der Analyse aufgrund fehlender Daten und die daraus resultierende mögliche Beeinflussung des Ergebnisses für seltene Ereignisse.

### **Fazit**

Es konnte auf der Basis von 2 randomisierten kontrollierten Studien zum Vergleich zwischen einer größeren und einer geringeren Erfahrung / Expertise, befördert durch verschiedene Weiterbildungsmaßnahmen, für insgesamt 5 der vordefinierten Endpunkte eine statistisch signifikante Auswirkung nachgewiesen werden. Die auf Schulungsmaßnahmen und Audit- und Feedbackstrategien basierenden Weiterbildungsstrategien mit dem Ziel, die Kaiserschnitttrate zu senken, führten bei initial geplanter vaginaler Geburt bei insgesamt niedriger Qualität der Evidenz zu einer abnehmenden neonatalen Mortalität und Veränderungen im Geburtsmodus. Die Anzahl vaginaler Geburten nahm zu, wobei nicht zwischen Spontangeburt und vaginal-operativen Geburten unterschieden wurde, während die Auswirkung auf die Anzahl der Kaiserschnitte unklar bleibt. Diese 2 Endpunkte wurden von der Leitliniengruppe als für die Entscheidung kritisch eingestuft. Zusätzlich wurde eine Auswirkung auf 3 klinisch wichtige, aber für die Empfehlung nicht entscheidende Endpunkte gezeigt. Der Einsatz von Wehenmitteln zur Einleitung der Geburt stieg an, während er unter der Geburt sank. Bei den Neugeborenen wurde häufiger eine fortgeschrittene metabolische Acidose diagnostiziert. Weiterbildungsmaßnahmen zeigten keinen Einfluss auf die Endpunkte schwere metabolische Acidose, Amniotomie und Episiotomie. Auswirkungen auf die maternale Mortalität, das Auftreten höhergradiger Dammriss und Infektionen bleiben unklar. Die Herabstufungen der Qualität der Evidenz erfolgten nach dem GRADE-Ansatz und basieren auf schwerwiegenden Studienlimitationen, Indirektheit und schwerwiegender fehlender Genauigkeit.

Für die systemimmanenten Faktoren konnte zum Vergleich zwischen einer größeren und einer geringeren Erfahrung / Expertise, gemessen an der Geburtenzahl pro Jahr, auf der Basis von 6 Beobachtungsstudien für insgesamt 3 der vordefinierten Endpunkte eine statistisch signifikante Auswirkung nachgewiesen werden. Eine größere Geburtenzahl führte bei insgesamt sehr niedriger Qualität der Evidenz zu einer Veränderung im Geburtsmodus mit mehr vaginalen Geburten, mehr vaginal-operativen Geburten und weniger ungeplanten Kaiserschnitten. Bei den Neugeborenen wurde in größeren Geburtskliniken bei mehr Neugeborenen eine schwere metabolische Acidose diagnostiziert. Diese 2 Endpunkte wurden von der Leitliniengruppe als für die Entscheidung kritisch eingestuft. Zusätzlich wurde eine Auswirkung auf 1 klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkten gezeigt. Gebärende mit einer PDA erhielten in größeren Geburtskliniken seltener Wehenmittel während der Geburt. Auswirkungen auf die neonatale Mortalität und das Auftreten von Episiotomien bleiben unklar. Die Herabstufungen der Qualität der Evidenz erfolgten nach dem GRADE-Ansatz und basieren vorrangig auf dem Studiendesign der Beobachtungsstudien, schwerwiegenden Studienlimitationen, schwerwiegender Indirektheit und schwerwiegender fehlender Genauigkeit.

Für die Bewertung von Erfahrung / Expertise anhand eines weiteren systemimmanenten Faktors, der Versorgungsstufe der Geburtskliniken, basiert die Evidenz auf 1 Beobachtungsstudie. Eine statistisch signifikante Auswirkung konnte für 1 der vordefinierten klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte nachgewiesen werden. Eine höhere Versorgungsstufe führte bei insgesamt sehr niedriger Qualität der Evidenz dazu, dass Gebärende mit einer PDA seltener Wehenmittel während der Geburt erhielten. Die Auswirkung auf den Geburtsmodus und die Häufigkeit von Episiotomien bleiben unklar. Die Herabstufungen der Qualität der Evidenz erfolgten nach dem GRADE-Ansatz und basieren vorrangig auf dem Studiendesign der Beobachtungsstudien, schwerwiegender Studienlimitationen und schwerwiegender fehlender Genauigkeit.

Insgesamt ist zu beachten, dass nur für 5 der vordefinierten kritischen Endpunkte und für 5 der wichtigen Endpunkte Evidenz vorlag und mehrere vordefinierte Endpunkte nicht untersucht werden konnten.