

IQWiG-Berichte – Nr. 705

**Evidenzbericht zur
Erfahrung / Expertise des
geburtshilflichen
Fachpersonals**

Rapid Report

Auftrag: V16-01B
Version: 1.0
Stand: 20.12.2018

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Evidenzbericht zur Erfahrung / Expertise des geburtshilflichen Fachpersonals

Auftraggeber:

Bundesministerium für Gesundheit

Datum des Auftrags:

14.09.2016

Interne Auftragsnummer:

V16-01B

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Im Mediapark 8

50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Dieser Bericht wurde ohne die Beteiligung externer Sachverständiger erstellt. Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Mitarbeiter des IQWiG

- Susanne Unverzagt
- Nina Peterwerth
- Katharina Hirsch
- Claudia Mischke
- Inga Overesch
- Katrin Dreck
- Wiebke Sieben

Schlagwörter: Termingerechte Geburt, Krankenpflegepersonen – Geburtshilfe-, Neonatalogen, Anästhesisten, Systematische Übersicht

Keywords: Term Birth, Nurse Midwives, Neonatologists, Anesthetists, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	viii
Abbildungsverzeichnis	xi
Abkürzungsverzeichnis	xii
Kurzfassung	xiii
1 Hintergrund	1
2 Fragestellung	3
3 Projektverlauf	4
4 Methoden	5
4.1 Kriterien für den Einschluss von Studien zur Erfahrung / Expertise des geburtshilflichen Fachpersonals in die Untersuchung	5
4.1.1 Population.....	5
4.1.2 Prüf- und Vergleichsintervention	5
4.1.3 Endpunkte.....	5
4.1.4 Studientypen	7
4.1.5 Studiendauer	8
4.1.6 Publikationszeitraum	8
4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	8
4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen.....	9
4.2 Umfassende Informationsbeschaffung	9
4.2.1 Primäre Informationsquellen	9
4.2.1.1 Bibliografische Datenbanken.....	9
4.2.1.2 Studienregister	10
4.2.1.3 Weitere Informationsquellen und Suchtechniken.....	10
4.2.1.3.1 Studiengruppen und Fachgesellschaften	10
4.2.1.3.2 Weitere Suchtechniken.....	10
4.2.1.4 Autorenanfragen	10
4.2.2 Selektion relevanter Studien.....	10
4.3 Informationsextraktion und -bewertung auf Studienebene	11
4.3.1 Datenextraktion	11
4.3.2 Bewertung des Verzerrungspotenzials der Ergebnisse.....	11
4.3.3 Bewertung des Evidenzniveaus aller eingeschlossenen Studien.....	13
4.4 Informationssynthese und -analyse	13
4.4.1 Gegenüberstellung der Ergebnisse der Einzelstudien	13

4.4.2	Metaanalysen.....	14
4.4.3	Aussagen zur Qualität der Evidenz	14
4.4.4	Sensitivitätsanalysen	17
4.4.5	Subgruppenmerkmale und andere Effektmodifikatoren	17
4.4.6	Evidenzprofil	18
5	Ergebnisse aus der umfassenden Informationsbeschaffung.....	20
5.1	Bibliografische Datenbanken.....	20
5.2	Studienregister	21
5.3	Weitere Informationsquellen und Suchtechniken	22
5.3.1.1	Studiengruppen und Fachgesellschaften.....	22
5.3.1.2	Anwendung weiterer Suchtechniken	22
5.3.1.3	Autorenanfragen	22
5.4	Resultierender Studienpool.....	22
5.5	Studien ohne berichtete Ergebnisse.....	23
6	Ergebnisse zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	24
6.1	Charakteristika der in die Bewertung eingeschlossenen Studien.....	24
6.1.1	Studiendesign und Studienpopulation	24
6.1.2	Übersicht der bewertungsrelevanten Endpunkte.....	34
6.2	Einschätzung des Verzerrungspotenzials	36
6.2.1	Einschätzung des Verzerrungspotenzials auf Studienebene	36
6.2.2	Einschätzung des Verzerrungspotenzials auf Endpunktebene	37
6.3	Ergebnisse zu kritischen Endpunkten.....	39
6.3.1	Maternale Mortalität.....	40
6.3.2	Geburtsmodus.....	40
6.3.3	Höhergradige Dammrisse (III°, IV°).....	43
6.3.4	Neonatale Mortalität.....	43
6.3.5	Apgar-Score nach 5 Minuten < 6	44
6.3.6	Schwere metabolische Acidose (pH < 7,0)	44
6.4	Ergebnisse zu wichtigen Endpunkten	44
6.4.1	Episiotomie.....	44
6.4.2	Wehenmittel	45
6.4.3	Amniotomie.....	45
6.4.4	Infektionen.....	45
6.4.5	Fortgeschrittene metabolische Acidose.....	45
7	Ergebnisse zur Erfahrung / Expertise durch systemimmanente Faktoren.....	47
7.1	Charakteristika der in die Bewertung eingeschlossenen Studien.....	47
7.1.1	Studiendesign und Studienpopulationen	47

7.1.2	Übersicht der bewertungsrelevanten Endpunkte	61
7.2	Einschätzung des Verzerrungspotenzials	63
7.2.1	Einschätzung des Verzerrungspotenzials auf Studienebene	63
7.2.2	Einschätzung des Verzerrungspotenzials auf Endpunktebene	65
7.3	Ergebnisse zu kritischen Endpunkten.....	66
7.3.1	Geburtsmodus.....	67
7.3.2	Neonatale Mortalität.....	70
7.3.3	Apgar-Score nach 5 Minuten < 6	71
7.3.4	Schwere metabolische Acidose (pH < 7,0)	71
7.3.5	Verlegung auf die Neugeborenen-Intensivstation	71
7.4	Ergebnisse zu wichtigen Endpunkten	72
7.4.1	Episiotomie.....	72
7.4.2	Wehenmittel	74
8	Evidenzprofil für die kritischen und wichtigen Endpunkte	76
8.1	Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	92
8.1.1	Maternale Mortalität.....	93
8.1.2	Geburtsmodus.....	93
8.1.3	Höhergradige Dammrisse (III°, IV°).....	94
8.1.4	Neonatale Mortalität.....	95
8.1.5	Schwere metabolische Acidose (pH < 7,0)	95
8.1.6	Episiotomie.....	96
8.1.7	Wehenmittel	96
8.1.8	Amniotomie.....	97
8.1.9	Infektionen.....	98
8.1.10	Fortgeschrittene metabolische Acidose.....	99
8.1.11	Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte.....	99
8.2	Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)	100
8.2.1	Geburtsmodus.....	101
8.2.2	Neonatale Mortalität.....	103
8.2.3	Schwere metabolische Acidose (pH < 7,0)	104
8.2.4	Episiotomie.....	105
8.2.5	Wehenmittel	105
8.2.6	Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte.....	106
8.3	Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe).....	107
8.3.1	Geburtsmodus.....	108

8.3.2	Episiotomie.....	108
8.3.3	Wehenmittel	109
8.3.4	Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte.....	110
9	Diskussion.....	111
10	Fazit.....	120
11	Literatur	122
12	Studienlisten	132
12.1	Liste der eingeschlossenen Studien.....	132
12.2	Liste der gesichteten systematischen Übersichten	133
12.3	Liste der ausgeschlossenen Publikationen zum Thema mit Ausschlussgründen.....	133
Anhang A – Suchstrategien		143
A.1	– Bibliografische Literaturrecherche	143
A.2	– Suche in Studienregistern	150
Anhang B – Verzerrungspotenzial auf Endpunktebene.....		152
B.1	– Erfahrung / Expertise durch Weiterbildungsmaßnahmen auf der Grundlage von randomisierten kontrollierten Studien	152
B.1.1	– Maternale Mortalität	152
B.1.2	– Geburtsmodus	153
B.1.3	– Höhergradige Dammrisse (III°, IV°)	153
B.1.4	– Neonatale Mortalität	154
B.1.5	– Schwere metabolische Acidose.....	154
B.1.6	– Episiotomie	155
B.1.7	– Wehenmittel.....	156
B.1.8	– Amniotomie	156
B.1.9	– Infektionen	157
B.1.10	– Leichte, mittelgradige und fortgeschrittene Acidose	157
B.2	– Erfahrung / Expertise durch systemimmanente Faktoren.....	158
B.2.1	– Geburtsmodus	158
B.2.2	– Neonatale Mortalität	158
B.2.3	– Schwere metabolische Acidose (pH < 7,0).....	159
B.2.4	– Episiotomie	159
B.2.5	– Wehenmittel	160
Anhang C – Ergebnisse.....		161
C.1	– Ergebnisse zu Erfahrung / Expertise durch Weiterbildungsmaßnahmen	161
C.1.1	– Maternale Mortalität	162
C.1.2	– Geburtsmodus	162
C.1.3	– Höhergradige Dammrisse (III°, IV°)	166

C.1.4 – Neonatale Mortalität	167
C.1.5 – Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6	169
C.1.6 – Schwere metabolische Acidose (pH < 7,0)	170
C.1.7 – Episiotomie	171
C.1.8 – Wehenmittel	172
C.1.9 – Amniotomie	173
C.1.10 – Infektionen	174
C.1.11 – Fortgeschrittene metabolische Acidose	175
C.2 – Ergebnisse zu Erfahrung / Expertise durch systemimmanente Faktoren	175
C.2.1 – Geburtsmodus	175
C.2.2 – Neonatale Mortalität	179
C.2.3 – Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6	180
C.2.4 – Schwere metabolische Acidose (pH < 7,0)	181
C.2.5 – Ergänzende Informationen zur Verlegung auf die Neugeborenen- Intensivstation	182
C.2.6 – Episiotomie	183
C.2.7 – Wehenmittel	185

Tabellenverzeichnis

	Seite
Tabelle 1: Durch die Leitliniengruppe konsentierete Endpunkte für vergleichende Interventionsstudien und ihre Bewertung	7
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss.....	9
Tabelle 3: Übersicht über Autorenanfragen	22
Tabelle 4: Studienpool der Evidenzbewertung	23
Tabelle 5: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse	23
Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	27
Tabelle 7: Matrix der kritischen und klinisch wichtigen Endpunkte zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	35
Tabelle 8: Bewertung des Verzerrungspotenzials auf Studienebene zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	37
Tabelle 9: Zusammenfassung der Ergebnisse zu den kritischen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	40
Tabelle 10: Zusammenfassung der Ergebnisse zu den wichtigen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	44
Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen	50
Tabelle 12: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von Geburtsregistern	57
Tabelle 13: Matrix der kritischen und klinisch wichtigen Endpunkte zur Erfahrung / Expertise durch systemimmanente Faktoren	62
Tabelle 14: Verzerrungspotenzial auf Studienebene zur Erfahrung / Expertise durch systemimmanente Faktoren.....	65
Tabelle 15: Zusammenfassung der Ergebnisse zu den kritischen Endpunkten zur Erfahrung / Expertise durch systemimmanente Faktoren	67
Tabelle 16: Zusammenfassung der Ergebnisse zu den wichtigen Endpunkten zur Erfahrung / Expertise durch systemimmanente Faktoren	72
Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	77
Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)	82
Tabelle 19: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe).....	89
Tabelle 20: Bewertung des Verzerrungspotenzials auf Endpunktebene zu maternale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	152

Tabelle 21: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	153
Tabelle 22: Bewertung des Verzerrungspotenzials auf Endpunktebene zu höhergradige Dammrisse (III°, IV°) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	153
Tabelle 23: Bewertung des Verzerrungspotenzials auf Endpunktebene zu neonataler Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	154
Tabelle 24: Bewertung des Verzerrungspotenzials auf Endpunktebene zu schwerer metabolischer Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	154
Tabelle 25: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Episiotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	155
Tabelle 26: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Wehenmitteln unter der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	156
Tabelle 27: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Wehenmittelgabe zur Einleitung der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	156
Tabelle 28: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Amniotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	157
Tabelle 29: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Infektion zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	157
Tabelle 30: Bewertung des Verzerrungspotenzials auf Endpunktebene zu fortgeschrittener Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	158
Tabelle 31: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Geburtsmodus zur Erfahrung / Expertise durch systemimmanente Faktoren	158
Tabelle 32: Bewertung des Verzerrungspotenzials auf Endpunktebene zu neonataler Mortalität zur Erfahrung / Expertise durch systemimmanente Faktoren	159
Tabelle 33: Bewertung des Verzerrungspotenzials auf Endpunktebene: schwere metabolische Acidose zur Erfahrung / Expertise durch systemimmanente Faktoren	159
Tabelle 34: Bewertung des Verzerrungspotenzials auf Endpunktebene: Episiotomie zur Erfahrung / Expertise durch systemimmanente Faktoren	160
Tabelle 35: Bewertung des Verzerrungspotenzials auf Endpunktebene: Wehenmittel zur Erfahrung / Expertise durch systemimmanente Faktoren	160
Tabelle 36: Ergebnisse für die maternale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	162
Tabelle 37: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (3 Ausprägungen des Endpunktes berichtet)	163
Tabelle 38: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (2 Ausprägungen des Endpunktes berichtet)	164
Tabelle 39: Ergebnisse für höhergradige Dammrisse (III°, IV°) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	166
Tabelle 40: Ergebnisse für die neonatale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	167

Tabelle 41: Ergänzende Informationen für die neonatale Mortalität (Totgeburten) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	168
Tabelle 42: Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6 (Apgar-Score nach 5 Minuten < 7) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	169
Tabelle 43: Ergebnisse für schwere metabolische Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	170
Tabelle 44: Ergebnisse für Episiotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	171
Tabelle 45: Ergebnisse für Wehenmittel während der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	172
Tabelle 46: Ergebnisse für Wehenmittel zur Einleitung der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	173
Tabelle 47: Ergebnisse für Amniotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	173
Tabelle 48: Ergebnisse für maternale Infektionen zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	174
Tabelle 49: Ergänzende Informationen für neonatale Infektionen zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen	174
Tabelle 50: Ergebnisse für die fortgeschrittene Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.....	175
Tabelle 51: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr der Geburtsklinik)	176
Tabelle 52: Ergebnisse zum Geburtsmodus zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe der Geburtskliniken).....	178
Tabelle 53: Ergebnisse zur neonatalen Mortalität zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr).....	179
Tabelle 54: Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6 (Apgar-Score nach 5 Minuten < 7 oder < 4 zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr))	180
Tabelle 55: Ergebnisse zur schweren metabolischen Acidose (pH < 7,0) zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)	181
Tabelle 56: Ergänzende Informationen für die Häufigkeit der Verlegung auf die Neugeborenen-Intensivstation zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr).....	182
Tabelle 57: Ergebnisse für die Häufigkeit von Episiotomien zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr).....	183
Tabelle 58: Ergebnisse für die Häufigkeit von Episiotomien zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe)	184
Tabelle 59: Ergebnisse für die Häufigkeit von Wehenmittelgaben während der Geburt zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr).....	186
Tabelle 60: Ergebnisse für die Häufigkeit der Wehenmittelgaben zur Einleitung der Geburt zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren	186

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion.....	21
Abbildung 2: Forest Plot für die vaginale Geburt.....	165
Abbildung 3: Forest Plot für den Kaiserschnitt.....	165
Abbildung 4: Forest Plot für Episiotomie zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)	185
Abbildung 5: Forest Plot für Episiotomie zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe).....	185

Abkürzungsverzeichnis

Abkürzung	Bedeutung
aOR	Adjustiertes Odds Ratio
Apgar-Score	Schema zur Beurteilung der postpartalen Adaptation und Vitalität des reifen Neugeborenen nach den Kriterien Atmung, Puls, Grundtonus, Aussehen und Reflexe
aRR	Adjustiertes relatives Risiko
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
BMG	Bundesministerium für Gesundheit
CONSORT	Consolidated Standards of Reporting Trials (konsolidierte Standards zum Berichten von Studien)
CTG	Kardiotokogramm
DAMI	Design Algorithm for Medical Literature on Intervention (Designklassifizierung für medizinische Interventionsstudien)
DGGG	Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V.
DGHWi	Deutsche Gesellschaft für Hebammenwissenschaft e. V.
EbM	Evidence-based Medicine (evidenzbasierte Medizin)
EP	Eröffnungsperiode
G-BA	Gemeinsamer Bundesausschuss
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation (Bewertung von Empfehlungen, Einschätzung, Entwicklung und Auswertung)
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KI	Konfidenzintervall
OCEBM	Oxford Centre for Evidence-based Medicine (Oxford-Zentrum für evidenzbasierte Medizin)
OR	Odds Ratio
PDA	Periduralanästhesie
PICO	Participants Intervention Comparison Outcomes (Akronym für Studienteilnehmer, Interventionsgruppe, Vergleichsgruppe, Endpunkte)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	Risikodifferenz
RR	relatives Risiko
SSW	Schwangerschaftswoche

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 ≥ 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 (pH $< 7,0$), Basendefizit ≥ -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 ≥ 1000 ml, Aufnahme einer Stillbeziehung bei initialem Stillwunsch, teilweise / vollständige Plazentaretention, Geburtserleben, subjektive Gesundheitseinschätzung, Apgar-Score nach 5 Minuten < 6, Basendefizit ≥ -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 Dammrisse (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 ≥ 1000 ml, zur Aufnahme einer Stillbeziehung bei initialem Stillwunsch, zur teilweisen / vollständigen Plazentaretention, zum Geburtserleben, zur subjektiven Gesundheitseinschätzung und zum Basendefizit ≥ -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.

1 Hintergrund

Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) wurde vom Bundesministerium für Gesundheit (BMG) beauftragt, die Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG) und die Deutsche Gesellschaft für Hebammenwissenschaft e. V. (DGHWi) bei der Entwicklung einer interdisziplinären S3-Leitlinie „Die vaginale Geburt am Termin“ (Registernummer der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften [AWMF] 015-083) [1] zu unterstützen.

Ziel dieser Leitlinie ist die Evaluierung evidenzbasierter Interventionen bei der vaginalen Geburt am Termin, die Geburtshelferinnen und Geburtshelfern mehr Sicherheit geben soll. Die meisten Schwangerschaften führen zu einer vaginalen Geburt ohne notwendige intensive geburtshilfliche Interventionen. Bei auftretenden Risiken ist eine intensivmedizinische Betreuung von Mutter und Kind essenziell. Daraus ergibt sich die Notwendigkeit einer risikoadaptierten Versorgungsplanung und -gestaltung.

In Deutschland lag die Mortalität für alle Neugeborenen gemessen an der Frühsterblichkeit (bis zum 7. Lebenstag verstorben) 2015 bei 186 je 100 000 Lebendgeborene und für die Spätmortalität (8. bis 28. Lebenstag verstorben) bei 42 je 100 000 Lebendgeborene [2]. Die Müttersterblichkeit lag bei rund 3 Frauen je 100 000 Lebendgeborene [3]. Als Risikofaktoren für Komplikationen rund um die Schwangerschaft und Geburt gelten ein höheres Alter der Mütter bei der 1. Geburt, Rauchen während der Schwangerschaft, eine geringe soziale Unterstützung, ein Migrationshintergrund, Adipositas und Mehrlingsschwangerschaften [4].

Der Anteil der Kaiserschnittentbindungen an allen Geburten in deutschen Krankenhäusern ist in den letzten Jahrzehnten von 17,3 % im Jahr 1994 auf 32,0 % im Jahr 2017 [5] angestiegen und zeigt eine hohe innerdeutsche Variabilität [4]. Bei 23,1 % aller Kaiserschnitte im Geburtsverlauf wird als Hauptindikation ein pathologisches Kardiotokogramm (CTG) oder eine protrahierte Geburt als Einzeldiagnose genannt, wobei ein pathologisches CTG sogar bei 41,2 % der durchgeführten ungeplanten Kaiserschnitte eine der dokumentierten Indikationen ist [5]. Hinzu kommen eine aufgrund forensischer Motive zunehmend veränderte Risikobewertung in der Schwangerenbetreuung und Geburtshilfe, veränderte Klinikorganisationen und eine abnehmende Erfahrung des geburtshilflichen Fachpersonals in der Betreuung komplizierter Geburtsverläufe [6].

Die interdisziplinäre S3-Leitlinie „Die vaginale Geburt am Termin“ soll die Evidenzbasierung der Begleitung der natürlichen (vaginalen) Geburt durch Hebammen und Entbindungspfleger und ärztliche Geburtshelferinnen oder Geburtshelfer verbessern. Das S3-Leitlinienniveau fordert eine systematische Recherche, Auswahl und Bewertung der vorliegenden wissenschaftlichen Evidenz zu den relevanten klinischen Fragestellungen und die Beschreibung des methodischen Vorgehens, sodass in der Leitlinie schließlich nachvollziehbar dargelegt werden kann, auf welcher Evidenz die gegebenen Empfehlungen basieren [7].

Ziel des geplanten Evidenzberichts ist es, für eine durch die Fachgesellschaften und die AWMF vorgegebene Schlüsselfrage eine Evidenzgrundlage zusammenzustellen und diese hinsichtlich ihrer Qualität zu bewerten. Diese Frage umfasst die Erfahrung / Expertise des geburtshilflichen Fachpersonals. Der Evidenzbericht soll von der Leitliniengruppe als Grundlage für die Entwicklung von evidenzbasierten Empfehlungen für die S3-Leitlinie herangezogen werden.

Dieser Rapid Report des IQWiG nutzt die Methoden der AWMF zum Aufbau von Evidenztabelle und der GRADE-Arbeitsgruppe (GRADE: Grading of Recommendations, Assessment, Development, and Evaluation) zur nachvollziehbaren Bewertung der Qualität der Evidenz.

2 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?

3 Projektverlauf

Das BMG hat am 14.09.2016 das IQWiG mit der Unterstützung der Entwicklung einer interdisziplinären S3-Leitlinie „Die vaginale Geburt am Termin“ beauftragt.

Auf Basis der Projektskizze wurde ein Rapid Report zur Evidenz zu den Auswirkungen von Erfahrung / Expertise des geburtshilflichen Fachpersonals erstellt. Dieser Bericht wurde an das BMG und die Leitliniengruppe der AWMF übermittelt und nach Abschluss aller Teilprojekte gemeinsam auf der Website des IQWiG veröffentlicht.

4 Methoden

Evidenzbewertung anhand von vergleichenden Interventionsstudien

Die Evidenz zur Fragestellung lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien unter Beachtung maternaler und neonataler Endpunkte bewerten. Idealerweise sollten Personen oder Krankenhäuser einer Gruppe randomisiert zugeteilt worden sein.

4.1 Kriterien für den Einschluss von Studien zur Erfahrung / Expertise des geburtshilflichen Fachpersonals in die Untersuchung

4.1.1 Population

Eingeschlossen wurden Studien, die Gebärende um den errechneten Geburtstermin ab der vollendeten 37. Schwangerschaftswoche (SSW) (37 + 0 SSW) bis zur vollendeten 42. SSW (42 + 0 SSW) von der aktiven EP bis zum Ende der Nachgeburtsperiode untersuchen. Die Nachgeburtsperiode umfasst die Zeit von der Geburt des Neugeborenen bis 2 Stunden nach der Plazentageburt [8].

4.1.2 Prüf- und Vergleichsintervention

In die Bewertung wurden Studien eingeschlossen, welche verschiedene Ausprägungen von Erfahrung und Expertise der ärztlichen Geburtshelferinnen oder Geburtshelfer (Ärztinnen, Ärzte, Fachärztinnen, Fachärzte, Assistentinnen und Assistenten), Hebammen und Entbindungspfleger in Krankenhäusern als Prüf- und Vergleichsintervention miteinander vergleichen. Erfahrung und Expertise kann durch folgende Faktoren beschrieben werden:

- Anstellungsverhältnisse (Vollzeit / Teilzeit / freiberuflich)
- Berufsjahre / Weiterbildungen
- Anzahl von Geburten pro Fachperson
- Registrierung / Qualifikation
- systemimmanente Faktoren
 - Vergleiche von Krankenhäusern mit unterschiedlichen Geburtenzahlen
 - Vergleiche von Krankenhäusern mit angeschlossener Kinderklinik, eigener Neonatologie oder Bereitschaft / Kooperationen mit entsprechenden Kliniken

In der Untersuchung des Einflusses systemimmanenter Faktoren werden keine Prüf- und Vergleichsinterventionen, sondern verschiedene Expositions- und Referenzgruppen miteinander verglichen.

4.1.3 Endpunkte

Die Bewertung der Evidenzgrundlage für die Prüfinderventionen erfolgte anhand von Endpunkten, die von der Leitliniengruppe nach den GRADE-Vorgaben konsentiert und hinsichtlich ihrer Bedeutung für die Gebärenden und Neugeborenen den Kategorien „kritisch

für die Entscheidung“ und „klinisch wichtig, aber für die Empfehlung nicht entscheidend“ zugeordnet wurden. Endpunkte, die aus der Perspektive der Gebärenden unbedingt in der Entscheidungsfindung für oder gegen eine Intervention berücksichtigt werden müssen, wurden der Kategorie „kritisch“ zugeordnet. Weitere Endpunkte, die ebenfalls berücksichtigt werden sollten, aber nicht unbedingt in die Entscheidungsfindung einbezogen werden müssen, wurden als „klinisch wichtig, aber für die Empfehlung nicht entscheidend“ kategorisiert [11,12] (siehe Tabelle 1). Als kritischer Endpunkt ergänzt wurde entsprechend den IQWiG-Methoden die Mortalität der Mutter oder des Neugeborenen. Diese Ergänzung wurde mit der Leitliniengruppe konsentiert. Es war geplant, dass alle Endpunkte mit validierten Messinstrumenten erfasst sein müssen.

Tabelle 1: Durch die Leitliniengruppe konsentrierte Endpunkte für vergleichende Interventionsstudien und ihre Bewertung

Rangfolge der Endpunkte	Endpunkte
Kritisch für die Entscheidung (Bewertung mit 7–9 Punkten durch die Leitliniengruppe)	<p>Maternale Endpunkte:</p> <ul style="list-style-type: none"> ▪ Mortalität ▪ Spontangeburt ▪ vaginal-operative Geburt ▪ Kaiserschnitt ▪ höhergradige Dammriss (III°, IV°) ▪ hoher mütterlicher Blutverlust (≥ 1000 ml) ▪ Aufnahme Stillbeziehung bei initialem Stillwunsch ▪ teilweise / vollständige Plazentaretention ▪ Geburtserleben ▪ subjektive Gesundheitseinschätzung <p>Neonatale Endpunkte:</p> <ul style="list-style-type: none"> ▪ Mortalität (Früh- und Spätsterblichkeit) ▪ Apgar-Score nach 5 Minuten < 6 ▪ schwere metabolische Acidose ($\text{pH} < 7,0$) ▪ Basendefizit ≥ -16 mmol/l ▪ Verlegung auf die Neugeborenen-Intensivstation
Klinisch wichtig, aber für die Empfehlung nicht entscheidend (Bewertung mit 4–6 Punkten durch die Leitliniengruppe)	<p>Maternale Endpunkte:</p> <ul style="list-style-type: none"> ▪ Episiotomie ▪ Wehenmittel ▪ Amniotomie ▪ Kristellerhilfe ▪ Infektionen <p>Neonatale Endpunkte:</p> <ul style="list-style-type: none"> ▪ leichte, mittelgradige, fortgeschrittene Acidose^a ▪ Ikterus ▪ Anämie
<p>a: Eine Acidose wird wie folgt kategorisiert: $\text{pH} 7,19-7,15$: leichte Acidose, $\text{pH} 7,14-7,10$: mittelgradige Acidose, $\text{pH} 7,09-7,0$: fortgeschrittene Acidose, $\text{pH} < 7,0$: schwere Acidose [13,14].</p> <p>Apgar-Score: Schema zur Beurteilung der postpartalen Adaptation und Vitalität des reifen Neugeborenen nach den Kriterien Atmung, Puls, Grundtonus, Aussehen und Reflexe</p>	

4.1.4 Studientypen

Es wurden prospektiv geplante vergleichende Studien (RCTs, nicht randomisierte kontrollierte Studien, Kohortenstudien und Fall-Kontroll-Studien) mit möglichst hohem Evidenzniveau (Evidence-based-Medicine(EbM)-Kriterien nach Oxford [15]) eingeschlossen, wobei die Designklassifizierung der Studien auf den Design-Algorithm-for-Medical-Literature-on-Intervention(DAMI)-Kriterien beruhen sollte [16].

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 der Evidenz einer medizinischen Intervention.

Da die auf RCTs basierende Datenlage nicht hinreichend war, wurden auch nicht randomisierte prospektiv geplante vergleichende Interventionsstudien mit zeitlich paralleler Vergleichsgruppe und adäquater Confounderkontrolle zur Evidenzbewertung herangezogen. Nicht randomisierte Studien ohne Confounderkontrolle oder mit einem Evidenzniveau geringer als 3 nach den EbM-Kriterien nach Oxford [15] sollten nicht berücksichtigt werden.

Für alle unter Abschnitt 4.1.2 genannten Interventionen und unter Abschnitt 4.1.3 genannten Endpunkte aus Interventionsstudien wäre eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar gewesen. Die Randomisierung könnte über eine Clusterrandomisierung der Praxen oder Krankenhäuser erfolgen.

4.1.5 Studiendauer

Eingeschlossen wurden Studien mit einer Mindestdauer, die den Zeitraum vom Beginn der aktiven EP bis zum Ende der Nachgeburtsperiode umfasste. Aufgrund von Unsicherheiten in den Definitionen der Geburtsphasen wurden geeignet erscheinende Studien im Zweifel eingeschlossen und der Zeitpunkt des Einschlusses der Teilnehmerinnen, wie in der Publikation berichtet, in die Evidenztabelle extrahiert.

4.1.6 Publikationszeitraum

Es wurden die unter Abschnitt 4.1.4 beschriebenen Studientypen mit einem Publikationsdatum ab 1990 eingeschlossen. Für die Einschränkung des Publikationszeitraumes wurde die Publikation von Rose und Schmied-Knittel aus dem Jahr 2011 [17] zugrunde gelegt, da diese die Umbruchphase von der „programmierten zur selbstbestimmten Geburt“ der 1980er-Jahre beschreibt. Diese Einschränkung soll zur Identifikation relevanter Literatur führen, die den heutigen geburtshilflichen Standard berücksichtigt und daher für die Erstellung der Leitlinienempfehlungen Relevanz besitzt.

4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In Tabelle 2 werden die Kriterien aufgelistet, die Studien erfüllen mussten, um für die Fragestellungen in die Bewertung eingeschlossen zu werden.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Population: alle Schwangeren am Termin unter der Geburt ab aktiver EP wie in Abschnitt 4.1.1 formuliert
E2	Prüfintervention / Exposition: größere Erfahrung / Expertise der Gynäkologin oder des Gynäkologen / der ärztlichen Geburtshelferinnen oder Geburtshelfer / Hebammen oder Entbindungspfleger wie in Abschnitt 4.1.2 formuliert
E3	Vergleichsintervention /Referenz: geringere Erfahrung / Expertise des unter E12 aufgelisteten Fachpersonals wie in Abschnitt 4.1.2 formuliert
E4	Erhebung kritischer oder wichtiger Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	Randomisierte kontrollierte Studien, ggf. nicht randomisierte prospektiv geplante vergleichende Studien mit möglichst hohem Evidenzniveau wie in 4.1.4 formuliert
E6	Vollpublikation ^a
E7	Publikation ab 1990 wie in 4.1.6 formuliert
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht, der den Kriterien des CONSORT- [18], TREND- [19] oder STROBE-Statements [20] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; EP: Eröffnungsperiode; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

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

Für die Einschlusskriterien E1 (Population), E2 (Interventionsgruppe) und E3 (Vergleichsgruppe) reichte es aus, wenn bei mindestens 80 % der eingeschlossenen Gebärenden dieses Kriterium erfüllt war. Es war geplant, für solche Studien auf entsprechende Subgruppenanalysen zurückzugreifen, aber solche Studien, bei denen die Einschlusskriterien bei weniger als 80 % erfüllt waren, nur dann einzuschließen, wenn entsprechende Subgruppenanalysen vorlagen.

4.2 Umfassende Informationsbeschaffung

4.2.1 Primäre Informationsquellen

4.2.1.1 Bibliografische Datenbanken

Die systematische Recherche nach relevanten Studien beziehungsweise Dokumenten wurde in folgenden bibliografischen Datenbanken durchgeführt:

- Suche nach Primärstudien in den Datenbanken MEDLINE, Embase, Cochrane Central Register of Controlled Trials und Cinahl

- Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE, Embase und Cinahl parallel zur Suche nach relevanter Primärliteratur sowie Suche in den Datenbanken Cochrane Database of Systematic Reviews und HTA Database

4.2.1.2 Studienregister

Die folgenden Studienregister wurden durchsucht:

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

4.2.1.3 Weitere Informationsquellen und Suchtechniken

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

4.2.1.3.1 Studiengruppen und Fachgesellschaften

Es erfolgte eine schriftliche Anfrage bei folgenden Fachgesellschaften und Studiengruppen:

- AWMF
- Cochrane Pregnancy and Childbirth Group

4.2.1.3.2 Weitere Suchtechniken

Systematische Übersichten wurden hinsichtlich weiterer relevanter Studien beziehungsweise Dokumente gesichtet.

4.2.1.4 Autorenanfragen

Es wurden Anfragen an Autorinnen und Autoren gestellt, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten ließen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen waren.

4.2.2 Selektion relevanter Studien

Selektion relevanter Studien beziehungsweise Dokumente aus den Ergebnissen der bibliografischen Recherche

Die durch die Suche in bibliografischen Datenbanken identifizierten und zu screenenden Treffer wurden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 2) bewertet. Als potenziell relevant erachtete Dokumente wurden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgten durch 2 Reviewerinnen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen den beiden Reviewerinnen aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Suchquellen

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

- Studienregister
- Informationen von Studiengruppen und Fachgesellschaften

Informationen aus den folgenden Suchquellen wurden von 1 Reviewerin auf Studien gesichtet, die diese dann in Bezug auf ihre Relevanz bewertete; eine 2. Reviewerin überprüfte den gesamten Prozess inklusive der Bewertungen:

- identifizierte systematische Übersichten

Sofern in einem der genannten Selektionsschritte Diskrepanzen auftraten, wurden diese jeweils durch Diskussion zwischen den beiden Reviewerinnen aufgelöst.

4.3 Informationsextraktion und -bewertung auf Studienebene

Alle für den Evidenzbericht relevanten Ergebnisse wurden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Genauigkeit der Ergebnisse, überprüft. Für jeden für die Entscheidung kritischen Endpunkt wurde eine endpunktspezifische studienübergreifende Aussage zur Qualität der Evidenz in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen und diese in eine der Kategorien hohe, moderate, niedrige und sehr niedrige Qualität der Evidenz entsprechend der GRADE-Guideline eingeteilt [21,22].

4.3.1 Datenextraktion

Alle für die Evidenzbewertung notwendigen Informationen wurden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Die Evidenztabellen orientierten sich an Vorlagen der AWMF.

Die Extraktion erfolgte durch 1 Reviewerin und wurde von einer 2. Reviewerin auf Grundlage der Primärstudien kontrolliert.

4.3.2 Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wurde für jede in die Evidenzbewertung eingeschlossene Studie bewertet, und zwar separat für jeden Endpunkt. Dazu wurden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet [23]:

A: Aspekte des Verzerrungspotenzials auf Studienebene

- Erzeugung der Randomisierungssequenz (für RCTs) oder zeitliche Parallelität der Gruppen (für nicht randomisierte Studien)

- Verdeckung der Gruppenzuteilung (für RCTs) oder Vergleichbarkeit der Gruppen / adäquate Berücksichtigung von prognostisch relevanten Faktoren (für nicht randomisierte Studien)
- Verblindung der Gebärenden sowie der behandelnden Person
- ergebnisgesteuerte Berichterstattung

Um den Vorgaben für Evidenzberichte der Leitliniengruppe der AWMF und der GRADE-Arbeitsgruppe zu entsprechen, wurden insbesondere auch folgende Aspekte bewertet:

- früher Abbruch bei hohen Behandlungseffekten
- Rekrutierungsunterschiede in den Clustern der clusterrandomisierten Studien und Vergleichbarkeit der Gruppen beziehungsweise Berücksichtigung prognostisch relevanter Faktoren (bei Cluster-RCTs) [24]
- fehlende / nicht adäquat angewandte Ein- und Ausschlusskriterien (v. a. beim Einschluss der Kontrollgruppe), Auswahl der exponierten und nicht exponierten Personen in Kohortenstudien aus unterschiedlichen Populationen (bei nicht randomisierten kontrollierten Studien)
- verzerrte Erhebung von Exposition und Outcome (uneinheitliche Messung der Exposition, uneinheitliche Endpunkterhebung in Abhängigkeit von der Exposition in Kohortenstudien) (bei nicht randomisierten kontrollierten Studien)

B: Aspekte des Verzerrungspotenzials auf Endpunktebene

- Verblindung der endpunkterhebenden Personen
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisgesteuerte Berichterstattung

Anhand dieser Aspekte wurde das Verzerrungspotenzial zusammenfassend als gering, mittel oder hoch eingestuft. Ein geringes Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Bei einem mittleren oder hohen Verzerrungspotenzial ist davon auszugehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern oder sogar stark verändern würden [23].

Für die Bewertung eines Endpunkts wurde zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter (A) aufgeführten Aspekte als gering, mittel oder hoch eingestuft. Das Verzerrungspotenzial wurde als gering beurteilt, wenn dieses für alle genannten Aspekte gering ist. Eine Einteilung in die Kategorie mittel oder hoch erfolgte, wenn es Einschränkungen in 1 oder mehreren Aspekten gab, die das Vertrauen in die Effektschätzungen herabsetzen oder sogar stark herabsetzen. Falls die Einstufung als mittel oder hoch erfolgte, wurde das

Verzerrungspotenzial für den Endpunkt auch als mittel oder hoch bewertet. Ansonsten fanden die unter (B) genannten endpunktspezifischen Aspekte Berücksichtigung.

4.3.3 Bewertung des Evidenzniveaus aller eingeschlossenen Studien

Zusätzlich erfolgte auf Wunsch der AWMF eine studienspezifische endpunktübergreifende Bewertung der Evidenz jeder eingeschlossenen Studie nach den Oxford-Centre-for-Evidence-based-Medicine(OCEBM)-Kriterien [25]. Diese Bewertung basiert auf dem Studiendesign und hat bei Studienlimitationen, einer fehlenden Genauigkeit der Effekte und Indirektheit eine Abwertung zur Folge. Zum Beispiel wurden RCTs mit hohem Verzerrungspotenzial vom Evidenzniveau 2 auf das Evidenzniveau 3 abgewertet.

4.4 Informationssynthese und -analyse

Die Informationen wurden für die kritischen Endpunkte einer Informationssynthese und -analyse unterzogen [26]. Da Evidenz für weniger als 7 kritische Endpunkte vorlag, wurde die Evidenz der klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte ebenfalls in die Informationssynthese eingeschlossen. Wenn möglich wurden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgte darüber hinaus in jedem Fall.

4.4.1 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse zu den in den Studien berichteten Endpunkten wurden im Bericht vergleichend beschrieben.

Es war geplant, in bestimmten Fällen auf eine Darstellung einzelner Ergebnisse bei unzureichender Erfassung des Endpunkts bei allen Gebärenden in der jeweiligen Studie beziehungsweise unvollständiger Berichterstattung zu verzichten. Dies sollte insbesondere zutreffen, wenn ein hoher Anteil der Gebärenden nicht in der Auswertung enthalten ist. Es war geplant, dass Ergebnisse in der Regel nicht in die Evidenzbewertung einfließen sollten, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Gebärenden basieren, das heißt, wenn der Anteil der Gebärenden, die nicht in der Auswertung berücksichtigt wurden, größer als 30 % war [27]. In der Literatur werden zum Teil bereits Auswertungen, in denen 20 % der Studienteilnehmenden nicht berücksichtigt werden, als nicht mehr aussagekräftig betrachtet [27]. Ausnahmen von dieser Regel waren zum Beispiel dann geplant, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke oder Cluster) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [28].

Ergebnisse sollten auch dann nicht in die Evidenzbewertung einbezogen werden, wenn der Unterschied in den Anteilen nicht berücksichtigter Gebärender zwischen der Interventions- und der Vergleichsgruppe größer als 15 Prozentpunkte war.

4.4.2 Metaanalysen

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren, sollten die Einzelergebnisse mithilfe von Metaanalysen quantitativ zusammengefasst werden. Für die statistische Auswertung wurden primär die Ergebnisse aus ITT-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Auswahl der Modelle für Metaanalysen erfolgte gemäß den Kriterien, die in den aktuellen Allgemeinen Methoden des IQWiG [29] genannt sind. Metaanalysen mit zufälligen Effekten sollten bei 5 oder mehr Studien primär mithilfe der Knapp-Hartung-Methode durchgeführt werden. Bei einer geringeren Anzahl von Studien sollte gegebenenfalls ein Modell mit festem Effekt berechnet oder eine qualitative Zusammenfassung vorgenommen werden. Falls die für eine Metaanalyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorlagen, wurden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

Für stetige Variablen war geplant, die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g , als Effektmaß einzusetzen. Bei binären Variablen sollten Metaanalysen primär anhand des relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen sollten auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen war geplant, ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und von den verfügbaren Daten zu verwenden [30].

Im Falle von mehrarmigen Studien war es möglich, dass mehrere Behandlungsarme für dieselbe Fragestellung von Interesse waren (z. B. unterschiedliche Schulungsarten). Sprach aus inhaltlichen bzw. klinischen Gründen nichts gegen die gemeinsame Betrachtung verschiedener Interventionen, wurden die Ergebnisse der entsprechenden Behandlungsarme zusammengefasst dem Kontrollarm gegenübergestellt. Auf diese Weise gingen die Ergebnisse des Kontrollarms auch in eine gegebenenfalls durchgeführte Metaanalyse nur einmal ein [31].

Die Effektschätzer und Konfidenzintervalle (KIs) aus den Studien wurden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgte die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität [32]. War die Heterogenität der Ergebnisse der Studien nicht bedeutsam, sollte der gemeinsame (gepoolte) Effekt inklusive KI dargestellt werden. Bei bedeutsamer Heterogenität sollte stattdessen das Prädiktionsintervall dargestellt und die Ergebnisse sollten nur in begründeten Ausnahmefällen gepoolt werden. Außerdem sollte untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu sollten methodische Faktoren (siehe Abschnitt 4.4.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.5), zählen.

4.4.3 Aussagen zur Qualität der Evidenz

Für jeden berichteten kritischen Endpunkt wurde nach den GRADE-Kriterien eine endpunktspezifische studienübergreifende Aussage zum Vertrauen in die Effektschätzungen

(Qualität der Evidenz) in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen und diese in eine der Kategorien hohe, moderate, niedrige und sehr niedrige Qualität der Evidenz eingeteilt [21,22]. In die Bewertung gingen das Design der eingeschlossenen Studien, Studienlimitationen, das Risiko von Publikationsbias, die Genauigkeit und Konsistenz der Effekte, die Übertragbarkeit und gegebenenfalls Aspekte zur Aufwertung der Qualität der Evidenz ein.

Bei einer hohen Qualität der Evidenz kann mit hoher Wahrscheinlichkeit davon ausgegangen werden, dass das Ergebnis der vorliegenden Studien nahe am wahren Interventionseffekt liegt. Diese Sicherheit nimmt mit geringerer Qualität der Evidenz ab.

Die Bewertung erfolgte durch 2 Reviewerinnen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen den beiden Reviewerinnen aufgelöst. Ergebnisse auf der Grundlage von RCTs starteten aufgrund ihres Designs bei einer hohen Qualität der Evidenz, Ergebnisse aus Beobachtungsstudien sollten bei einer niedrigen Qualität beginnen [21,22]. Diese Qualität konnte bei einem hohen Verzerrungspotenzial, einem hohen Risiko von Publikationsbias, einer fehlenden Genauigkeit der Effekte, Inkonsistenz oder einer eingeschränkten Übertragbarkeit abgewertet. Sehr große Effekte, eine Dosis-Wirkungs-Beziehung oder wenn die Berücksichtigung aller wesentlichen Confounder den Effekt konservativ beeinflussen würden, konnten zur einer Aufwertung der Qualität führen. Alle vorgenommenen Ab- oder Aufwertungen der Qualität der Evidenz wurden durch Fußnoten gekennzeichnet.

Im GRADE-Ansatz basiert die Stärke von Leitlinienempfehlungen auf dem Vertrauen, dass „die erwünschten Effekte einer Behandlungsstrategie die unerwünschten Effekte insgesamt überwiegen“ [33]. Daher war eine Gesamtbeurteilung des Vertrauens in die vorliegenden Effektschätzungen notwendig. Diese ergab sich aus der niedrigsten Qualität der Evidenz der kritischen Endpunkte [34,35].

A: Abwertung bei Studienlimitationen (Verzerrungspotenzial)

Die Qualität der Evidenz wurde aufgrund von schwerwiegenden oder sehr schwerwiegenden Limitationen um 1 oder 2 Kategorien herabgestuft, wenn starke Limitierungen in 1 oder mehreren der unter Abschnitt 4.3.2 genannten Kriterien das Vertrauen in die Effektschätzungen beeinträchtigen. Die Qualität der Evidenz wurde beispielsweise um 1 Kategorie herabgestuft, wenn die Effektschätzungen hauptsächlich auf Studien mit mittlerem Verzerrungspotenzial basieren oder wenn in einer Studie ein hohes Verzerrungspotenzial für einen der unter 4.3.2 genannten Aspekte oder ein mittleres Verzerrungspotenzial für mehrere Aspekte das Vertrauen in die Effektschätzungen beeinträchtigt. Die Qualität der Evidenz wurde beispielsweise um 2 Kategorien herabgestuft, wenn die Effektschätzungen hauptsächlich auf Studien mit hohem Verzerrungspotenzial basieren oder wenn in einer Studie ein hohes Verzerrungspotenzial für 1 oder mehrere Aspekte das Vertrauen in die Effektschätzungen stark beeinträchtigt [23,36].

B: Abwertung bei Publikationsbias

Es war geplant, die Qualität der Evidenz um 1 Kategorie herabzustufen, wenn ein Evidenzkörper ein hohes Risiko für einen Publikationsbias aufweist. Kriterien für die Annahme eines Publikationsbias umfassten einen Evidenzkörper, der vorrangig auf kleinen, kommerziell gesponserten Studien beruht, einen beobachteten Zusammenhang zwischen den Effektschätzungen und der Studiengröße (oder der Genauigkeit der Effekte) oder relevante Unterschiede in den Ergebnissen veröffentlichter und nicht veröffentlichter Studien zeigt. Weitere Kriterien sind Studienabbrüche auf der Grundlage von hohen Effektschätzungen oder Ergebnisse, die auf kleinen Fallzahlen und wenigen Ereignissen basieren [37,38].

C: Abwertung bei fehlender Genauigkeit der Effekte

Die Qualität der Evidenz wurde wegen schwerwiegender oder sehr schwerwiegender Ungenauigkeiten der Effekte um 1 oder 2 Kategorien herabgestuft. Als primäres Kriterium wurde hierfür das KI herangezogen. Unterschieden sich die klinischen Schlussfolgerungen aus der oberen beziehungsweise unteren Grenze des KI (stark), wurde eine Abwertung wegen (sehr) schwerwiegender fehlender Genauigkeit in Betracht gezogen. Bei kleinen Studien mit großen Effekten auf der Grundlage kleiner Ereigniszahlen konnte die Unterschreitung der Fallzahl für eine adäquat gepowerte Studie („optimal information size“) die Ursache einer Abwertung der Effektstärke wegen schwerwiegender Ungenauigkeiten sein [39,40].

D: Abwertung bei inkonsistenten (heterogenen) Effekten (Inkonsistenz)

Die Qualität der Evidenz sollte bei schwerwiegender unerklärter Heterogenität (Inkonsistenz) zwischen den Studienergebnissen um 1 oder 2 Kategorien herabgestuft werden. Kriterien für die Einschätzung einer möglichen Heterogenität der Studienergebnisse konnten eine große Varianz und unterschiedliche Richtungen der Effektschätzungen der Einzelstudien, eine geringe oder fehlende Überlappung der KIs und statistische Kriterien für eine bedeutsame Heterogenität wie ein geringer p-Wert oder ein hoher I²-Wert sein (Abschnitt 4.4.2). Bei einer hohen, durch die in Abschnitt 4.4.5 definierten Subgruppen erklärten Heterogenität sollte das Poolen der Effekte auf die betreffenden Subgruppen beschränkt werden [41,42].

E: Abwertung bei eingeschränkter Übertragbarkeit (Indirektheit)

Die Qualität der Evidenz wurde bei schwerwiegenden oder sehr schwerwiegenden Einschränkungen in der Übertragbarkeit auf die Zielpopulation um 1 oder 2 Kategorien herabgestuft. Eine eingeschränkte Übertragbarkeit konnte auf Abweichungen zwischen den PICO-Kriterien der Studien und den definierten Einschlusskriterien basieren. Beispielsweise wurden Abweichungen hinsichtlich des Zeitpunkts des Einschlusses der Gebärenden in die Studien oder der Versorgungsverhältnisse und der Qualifikationen der Hebammen beziehungsweise Entbindungspfleger zwischen Deutschland und den Ländern, in denen die Studien durchgeführt wurden, erwartet [43,44].

F: Aufwertung

Die Qualität der Evidenz sollte in methodisch guten Beobachtungsstudien und nicht randomisierten Interventionsstudien bei großen beziehungsweise sehr großen Effekten um 1 oder 2 Kategorien heraufgestuft werden. Eine Heraufstufung um 1 Kategorie war bei großen Effekten geplant, wenn der Punktschätzer für das relative Risiko zwischen 2 und 5 beziehungsweise zwischen 0,2 und 0,5 liegt, ohne dass plausible Confounder diesen Effekt verringern könnten. Eine Heraufstufung um 1 weitere Kategorie sollte bei sehr großen Effekten erfolgen, wenn das relative Risiko über 5 beziehungsweise unter 0,2 liegt, keine relevanten Verzerrungsrisiken vorliegen und die Effekte präzise (mit hinreichend geringen KIs) geschätzt werden können. Eine weitere Bedingung für eine Heraufstufung der Qualität der Evidenz war eine Dosis-Wirkungs-Beziehung oder wenn alle weiteren plausiblen Confounder oder Verzerrungen den gezeigten Effekt reduzieren oder, wenn kein Effekt beobachtet wurde, einen Effekt nahelegen würden [45,46].

4.4.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse waren Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bildeten sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen wie der Auswahl der Expositions- und Referenzgruppen. Derartige Sensitivitätsanalysen erfolgten unabhängig von gegebenenfalls weiteren Analysen, mit denen die Ergebnissicherheit eines beobachteten Effekts bewertet wird.

Es wurde angenommen, dass das Ergebnis solcher Sensitivitätsanalysen die für einen beobachteten Effekt abgeleiteten Aussagen zur Qualität der Evidenz beeinflussen kann. Ein als nicht robust eingestuftes Ergebnis kann laut den GRADE-Guidelines dazu führen, dass die Qualität der Evidenz für den betrachteten Endpunkt aufgrund von Inkonsistenz um 1 oder 2 Kategorien abgewertet wird [26,41].

4.4.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies konnten direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Interventionen sein. Im Gegensatz zu den in Abschnitt 4.4.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen bestand hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte sollte die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung sein. In die Untersuchung von Effektmodifikatoren sollten die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen werden. Außerdem sollten eigene Analysen in Form von Meta-regressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren erfolgen. Es war geplant, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Alter der Gebärenden
- Parität (Erst- vs. Mehrgebärende)
- Einschluss von Schwangeren mit besonderen Risiken (z. B. Zustand nach traumatischem Geburtserlebnis, vorheriger Kaiserschnitt, Mehrlinge, psychosoziale Belastung, Adipositas der Gebärenden, Migrationshintergrund)

Es war geplant, sofern sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben hätten, diese ebenfalls begründet einzubeziehen. Bei Identifizierung möglicher Effektmodifikatoren erfolgte eine Präzisierung der für die beobachteten Effekte abgeleiteten Aussagen zur Qualität der Evidenz. Beispielsweise konnten Ergebnisse für die Subgruppen getrennt berechnet und die Qualität der Evidenz für spezielle Subgruppen von Gebärenden eingeschränkt beurteilt werden.

4.4.6 Evidenzprofil

Für die Fragestellung wurde die Evidenz für alle kritischen Endpunkte aus allen vorliegenden Studien in einem Evidenzprofil zusammengefasst [26,47-49]. Dieses enthält für die Fragestellung und gegebenenfalls die in Abschnitt 4.4.5 genannten Subgruppen folgende Informationen:

- die Anzahl der Studien und der Designs
- eine detaillierte Bewertung der Qualität der Evidenz zu den Punkten Studienlimitationen, Inkonsistenz, Indirektheit, fehlende Genauigkeit und anderen Faktoren einschließlich des Publikationsbias
- die Summe der Anzahl der Ereignisse und der eingeschlossenen Gebärenden in den Interventions- und Vergleichsgruppen aller eingeschlossenen Studien
- die (relativen) Effektschätzungen für die beobachteten Interventionen mit ihren jeweiligen KIs
- der Unterschied zwischen den Risiken mit und ohne Intervention [50] (und die entsprechenden KIs). Dieser basiert auf dem Populationsrisiko und der relativen Effektschätzung und ihrem KI. Die typische Häufigkeit des Endpunkts in der deutschen Bevölkerung (Populationsrisiko) basiert auf repräsentativen Daten der stationären Versorgung in Deutschland [5]. Bei dichotomen Endpunkten wird die Ausprägung, in Abhängigkeit von der Häufigkeit der beschriebenen Ereignisse, auf 100 Geburten bei einem Populationsrisiko über 10 %, auf 1000 bei einem Populationsrisiko unter 10 % oder für sehr seltene Ereignisse auf 10 000 Geburten bezogen. Das Ziel war die Angabe ganzer Zahlen.
- die Einteilung der Qualität der Evidenz nach den in Abschnitt 4.4.3 genannten GRADE-Kriterien in 1 der Kategorien hoch, moderat, niedrig und sehr niedrig. Die Gründe der Auf- oder Abwertung werden in Fußnoten nachvollziehbar dargestellt [34,51].

- Kommentare wie relevant abweichende Ergebnisse aus den in Abschnitt 4.4.4 genannten Sensitivitätsanalysen, die in Fußnoten nachvollziehbar dargestellt werden sollten

Untersuchte eine der eingeschlossenen Studien nicht alle zuvor festgelegten Endpunkte, wurde dies in den Tabellen beschrieben. Es war geplant, die Ergebnisse aus Sensitivitätsanalysen zu nennen, wenn sie zu stark abweichenden Effektschätzungen führen. Wenn Metaanalysen nicht möglich waren, wurde die vorhandene Evidenz für diese Endpunkte dargestellt.

5 Ergebnisse aus der umfassenden Informationsbeschaffung

5.1 Bibliografische Datenbanken

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

Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A.1. Die letzte Suche fand am 01.10.2018 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt 12.3.

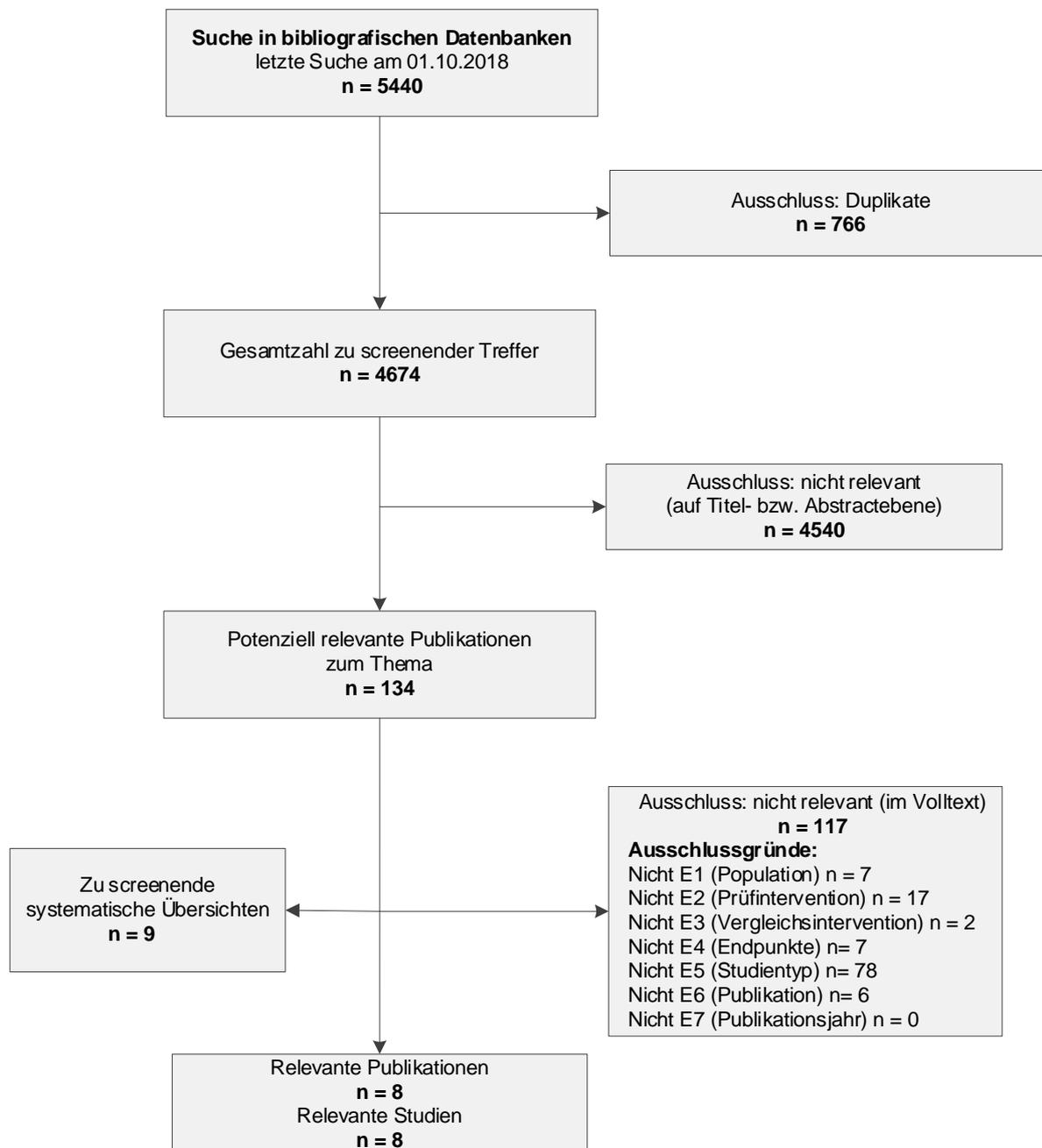


Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion

5.2 Studienregister

Durch die Suche in Studienregistern wurden keine relevanten Studien beziehungsweise Dokumente identifiziert. In den Studienregistern wurde 1 laufende Studie ohne berichtete Ergebnisse identifiziert (siehe Abschnitt 5.5).

Die Suchstrategien für die Suche in Studienregistern finden sich in Anhang A.2. Die letzte Suche in Studienregistern fand am 26.03.2018 statt.

5.3 Weitere Informationsquellen und Suchtechniken

Über weitere Informationsquellen und Suchtechniken identifizierte relevante Studien beziehungsweise Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Informationsquellen gefunden wurden.

5.3.1.1 Studiengruppen und Fachgesellschaften

Die schriftlich erfolgten Anfragen bei den Fachgesellschaften und Studiengruppen ergaben keine relevanten Studien beziehungsweise Dokumente, die nicht über andere Rechenschritte identifiziert werden konnten.

5.3.1.2 Anwendung weiterer Suchtechniken

Im Rahmen der Informationsbeschaffung wurden systematische Übersichten identifiziert – die entsprechenden Referenzen finden sich in Abschnitt 12.2. Die Referenzlisten dieser systematischen Übersichten wurden gesichtet.

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

5.3.1.3 Autorenanfragen

Für die vorliegende Evidenzbewertung wurden Autorenanfragen versendet (siehe Tabelle 3). Die Informationen aus den eingegangenen Antworten sind in die Studienbewertung eingeflossen.

Tabelle 3: Übersicht über Autorenanfragen

Studie / Studienregister-ID	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
Karalis 2016	Frage, ob die Registerdatenauswertung prospektiv geplant war und das Studienprotokoll zur Verfügung gestellt werden kann.	ja	Forschungsvorhaben war prospektiv geplant. Das Studienprotokoll kann nicht zur Verfügung gestellt werden, da es in finnischer Sprache verfasst wurde und nicht öffentlich ist.

5.4 Resultierender Studienpool

Durch die verschiedenen Suchschritte konnten insgesamt 8 relevante Studien (8 Dokumente) identifiziert werden (siehe Tabelle 4). Die entsprechenden Referenzen finden sich in Abschnitt 12.1.

Tabelle 4: Studienpool der Evidenzbewertung

Studie	Verfügbare Dokumente	
	Vollpublikation (in öffentlich zugänglichen Fachzeitschriften)	Ergebnisbericht aus Studienregistern beziehungsweise Eintrag Studienregister
Belghiti 2013	ja [52]	nein
Chaillet 2015	ja [53]	nein
Chuilon 2016	ja [54]	nein
Coulm 2012	ja [55]	nein
Finnström 2006	ja [56]	nein
Karalis 2016	ja [57]	nein
Lomas 1991	ja [58]	nein
Wehberg 2018	ja [59]	nein

5.5 Studien ohne berichtete Ergebnisse

In Tabelle 5 sind alle durch die Informationsbeschaffung identifizierten Studien ohne bisher berichtete Ergebnisse dargestellt.

Tabelle 5: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse

Studie	Dokumentart, ggf. Studienregister-ID, Zitat	Studientyp	Geplante Fallzahl	Status (ggf. geplantes Studienende)
OASI Care Bundle	Studienregistereintrag, ISRCTN12143325 [60]	Multizentrische Stepped-wedge trial	20 000	laufend (01.04.2019)

6 Ergebnisse zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

6.1 Charakteristika der in die Bewertung eingeschlossenen Studien

6.1.1 Studiendesign und Studienpopulation

Es wurden 2 RCTs zur Bewertung der Erfahrung / der Expertise durch Weiterbildungsmaßnahmen des geburtshilflichen Fachpersonals eingeschlossen (Chaillet 2015, Lomas 1991). Die Charakteristika sind in Tabelle 6 dargestellt.

In 1 Studie (Chaillet 2015) wurden 32 Krankenhäuser in Quebec in Kanada eingeschlossen. Im Durchführungszeitraum von 2008 bis 2011 wurden Daten zu insgesamt 184 952 Frauen erhoben. Es erfolgte eine Auswertung der Ergebnisse von 23 484 Erst- und Mehrgebärenden beziehungsweise 23 902 Neugeborenen in der Gruppe mit der Prüflintervention und 28 781 Erst- und Mehrgebärenden beziehungsweise 29 211 Neugeborene in der Gruppe mit der Vergleichsintervention. Das Ziel der Studie war die Senkung der Kaiserschnitttrate durch die Durchführung einer einjährigen Schulungsintervention des geburtshilflichen Personals in Audit und Feedbackmaßnahmen. In 16 Krankenhäusern beinhaltete die Prüflintervention die Schulung des geburtshilflichen Personals mit dem Programm Quality of Care, Obstetrics Risk Management, and Mode of Delivery (QUARISMA) und in 16 Krankenhäusern wurde als Vergleichsintervention keine Weiterbildungsmaßnahme für das geburtshilfliche Personal durchgeführt.

Eingeschlossen wurden kommunale, regionale und tertiäre, teilweise akademische Krankenhäuser mit ≥ 300 Geburten und einer Kaiserschnitttrate $> 17\%$ im Jahr vor der Initiierung der Studie und ohne kürzlich erfolgte oder laufende Qualitätsverbesserungsprogramme zur Senkung der Kaiserschnitttrate. Es wurden Gebärende ab der 24 + 0 SSW eingeschlossen. Über die eingeschlossenen Teilnehmerinnen liegen Charakteristika nur für die Gebärenden in der Präinterventionsphase vor. Charakteristika für Frauen in der Postinterventionsphase im Jahr nach der Intervention wurden nicht berichtet. Da jedoch im Jahr vor Einführung der Intervention $> 90\%$ der Gebärenden nach der 37 + 0 SSW geboren, ist davon auszugehen, dass das Kriterium, dass sich mindestens 80 % der eingeschlossenen Gebärenden in oder über der 37 + 0 SSW befinden mussten, nicht verletzt wurde. Ausschlusskriterien wurden in der Studie nicht beschrieben.

Das QUARISMA-Programm wurde in einem 1,5-jährigen Interventionszeitraum eingeführt. In den ersten 6 Monaten dieses Zeitraumes erfolgte eine Schulung zur evidenzbasierten klinischen Praxis durch Dozenten der Society of Obstetricians and Gynecologists of Canada, das geburtshilfliche Personal der jeweiligen Krankenhäuser wählte eine meinungsführende Person und benannte einen Auditausschuss. Es wurden 1-Tages-Workshops zur Durchführung von Audit- und Feedback-Maßnahmen, zur Verbesserung der Indikationsstellung eines intrapartalen Kaiserschnitts sowie zur Verbesserung des intrapartalen Managements durchgeführt. Im Jahr nach dieser Trainingsphase erfolgten 4 durch den Auditausschuss organisierte Auditzyklen mit externer Unterstützung über jeweils 3 Monate. Während der folgenden 1-

jährigen Postinterventionsphase wurde das Gesundheitsfachpersonal ermutigt, die klinischen Audits auch ohne Supervision weiterzuführen.

In den übrigen 16 Krankenhäusern der Vergleichsintervention wurden keine Maßnahmen durchgeführt. Zur Vermeidung einer Effektverdünnung wurden alle Programme zur Qualitätsverbesserung jährlich überprüft.

In Lomas 1991 wurden 16 Krankenhäuser in Ontario in Kanada eingeschlossen und im Durchführungszeitraum von 1988 bis 1989 Ergebnisse zu insgesamt 2496 Frauen mit einem vorherigen Kaiserschnitt ausgewertet, von denen bei 914 Frauen eine vaginale Geburt angestrebt wurde. Das Ziel der 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 Prüfindervention eine Schulung von Meinungsführern 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.

Eingeschlossen wurden Krankenhäuser ohne den Status eines Lehrinstitutes mit mindestens 100 Betten, wovon mindestens 10 geburtshilflich sein mussten. Es wurden Gebärende mit 1 vorherigen niedrigen transversalen Kaiserschnitt, einer Einlingsschwangerschaft in Schädel-lage und ohne absolute Indikationen für einen Kaiserschnitt eingeschlossen. Ausgeschlossen wurden Krankenhäuser, die in einem Landkreis mit einer Lehrinstitution oder im Norden Ontarios lagen, sowie Gebärende mit mehr als einem vorherigen Kaiserschnitt, einer vertikalen uterinen Narbe oder sonstigen Kontraindikationen entsprechend den Leitlinienkriterien. Es liegen keine Informationen zum Gestationsalter der eingeschlossenen Gebärenden vor. Ausgehend von einer normalen Schwangerschaftsdauer von 40 Wochen (40 + 0 SSW) und einer normalverteilten Abweichung vom Geburtstermin, nach der nur 17,5 % aller Frauen mehr als 10 Tage vor dem errechneten Geburtstermin gebären [8], wurde davon ausgegangen, dass mindestens 80 % der Frauen nach der vollendeten 37. SSW gebären.

Die eingeschlossene Studie untersuchte 2 Strategien zur Umsetzung einer Leitlinie zum Geburtsmanagement bei Frauen mit einem vorherigen Kaiserschnitt. In den Krankenhäusern, die der Weiterbildungsmaßnahme über spezifisch geschulte ärztliche Geburtshelfer („Schulung von Meinungsführern“) zugewiesen wurden, wurde vom geburtshilflichen ärztlichen Personal jeweils ein Meinungsführer gewählt, welcher an einem 1,5-tägigen Workshop zur Evidenz der Empfehlungen der Leitlinie und zu Grundprinzipien der Verhaltensänderung teilnehmen sollte. Es wurden mehrere Informationsschreiben im Namen des jeweiligen Meinungsführers mit Informationen zur Leitlinie an das geburtshilflich tätige ärztliche Personal des jeweiligen Krankenhauses versandt. Der jeweilige Meinungsführer sollte eine Veranstaltung mit einem fachkundigen, kompetenten und glaubwürdigen Referenten im Bereich der vaginalen Geburt nach vorherigem Kaiserschnitt organisieren und den regelmäßigen formellen sowie informellen Kontakt mit Kolleginnen und Kollegen aufrechterhalten und ausbauen. Alle Meinungsführer

waren männliche ärztliche Geburtshelfer, 3 der 4 Meinungsführer hatten eine klare mehrheitliche Unterstützung von ihren Kolleginnen und Kollegen.

In der Audit-und-Feedback-Gruppe wurden nach Festlegung von Kriterien für die Durchführung eines Kaiserschnitts bei allen Frauen mit einem vorherigen Kaiserschnitt die Dokumentationsakten überprüft, und es wurde die tatsächliche Praxis im 1. Jahr der Studiedurchführung mit den vereinbarten Kriterien verglichen. Alle 3 Monate wurden die Audit-Ergebnisse und insbesondere Diskrepanzen in der Geburtsklinik mit allen Geburtshelfern diskutiert. In 2 Krankenhäusern organisierten die Ärzte eigenständig die Überprüfung der Dokumentationsakten, in den anderen beiden Krankenhäusern führte das Studienteam die Überprüfung durch. Das Feedback wurde vom Studienteam gegeben, das bei der Abteilungsbesprechung die jeweiligen Ergebnisse und insbesondere die Fälle, für die keine dokumentierte Begründung für eine Abweichung von den Kriterien gefunden wurde, präsentierte.

In den übrigen 8 Krankenhäusern der Vergleichsintervention wurde lediglich der Verbreitungsprozess, der auf die ursprüngliche Veröffentlichung der Leitlinie folgte, wiederholt. Zu Beginn der Studie wurde allen in der Geburtshilfe tätigen Personen eine Kopie der Leitlinie zugesandt. Ein kurzes Anschreiben fokussierte auf den Teil der Leitlinie, der sich mit der Anwendung des Kaiserschnitts bei Frauen mit einem vorherigen Kaiserschnitt befasste. Darüber hinaus wurden keine weiteren Interventionen vorgenommen.

Bezüglich der Umsetzung der beiden Weiterbildungsmaßnahmen beschreiben die Autorinnen und Autoren, dass die Wahl des Meinungsführers auf einer durchschnittlichen Rücklaufquote der Fragebogen zur Identifizierung des Meinungsführers von 65,3 % basierte und Informationen, welcher Anteil der ärztlichen Geburtshelfer die anschließend festgelegten Meinungsführer wählten, nicht berichtet wurden. 3 der Meinungsführer führten die vorgesehene Veranstaltung mit einem Fachreferenten durch. Im 4. Krankenhaus schien der Meinungsführer nur 1 von 2 Fraktionen zu vertreten, er beteiligte sich nach den ersten Monaten nicht mehr an den weiteren Aktivitäten und nach 6 Monaten trat er von seiner Teilnahme zurück. Die Aktivitäten der Meinungsführer werden bezogen auf Art und Umfang der Informationsvermittlung als heterogen beschrieben. Alle geplanten Audit- und Feedback-Besprechungen wurden in 2 der 4 Krankenhäuser durchgeführt. In den beiden anderen Krankenhäusern fanden 3 der 4 Treffen statt und das Personal erhielt das Feedback zum nicht stattgefundenen Treffen per Post. Das Feedback erfolgte ausschließlich auf Abteilungsebene, ohne die Inanspruchnahme eines individuellen Feedbacks des ärztlichen Personals.

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungspotenzial auf Studienebene ^d	Evidenzniveau nach OCEBM
<p>Chaillet 2015 Cluster-RCT</p> <p>Quebec, Kanada 2008–2011</p> <p>Evaluation des Einflusses einer Schulungsstrategie auf die Kaiserschnitttrate</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden <ul style="list-style-type: none"> ▫ Krankenhäuser mit ≥ 300 Geburten in dem Jahr vor Initiierung der Studie, einer Kaiserschnitttrate > 17 % und ohne zum Zeitpunkt der Rekrutierung kürzlich erfolgten oder laufenden Qualitätsverbesserungsprogrammen zur Senkung der Kaiserschnitttrate. ▫ alle Gebärenden der teilnehmenden Krankenhäuser mit einem Gestationsalter ≥ 24+0 SSW und einem Neugeborenen mit einem Geburtsgewicht ≥ 500 g. ▪ Es wurden keine Ausschlusskriterien formuliert. 	<ul style="list-style-type: none"> ▪ 32 Krankenhäuser: 16 vs. 16 <ul style="list-style-type: none"> ▫ Art des Krankenhauses: <ul style="list-style-type: none"> - Kommunales Krankenhaus: 5,4 % vs. 2,8 % - Regionales Krankenhaus: 65,8 % vs. 65,1 % - Tertiäres Krankenhaus: 28,8 % vs. 32,1 % - Akademisches Krankenhaus: 36,8 % vs. 56,3 % ▪ 184 952 Gebärende insgesamt: <ul style="list-style-type: none"> ▫ 24 388 vs. 28 698 Gebärende in der Prä-interventionsphase <ul style="list-style-type: none"> - Alter^e: 29,4 vs. 29,8 - Parität: 44,0 % vs. 45,9 % Nullipara, Erstgebärende 36,5 % vs. 37,0 %, Mehrgebärende 19,6 % vs. 49,26 % - 37–41 SSW: 91,3 % vs. 93,0 % 	<ul style="list-style-type: none"> ▪ Schulung des geburtshilflichen Personals in Audit- und Feedback-Maßnahmen (QUARISMA-Programm) über 1,5 Jahre (n = 16 Krankenhäuser, 84 227 Gebärende, davon 23 484 in der Postinterventionsphase) <ul style="list-style-type: none"> ▫ Trainingsphase (6 Monate): <ul style="list-style-type: none"> - erste Schulung in evidenzbasierter klinischer Praxis vor Ort durch Dozenten der Society of Obstetricians and Gynecologists of Canada - Benennung eines Meinungsführers - Durchführung von Audits und Feedback (1-Tages-Workshop) - Verbesserung der Indikationsstellung für einen Kaiserschnitt und des intrapartalen Managements (1-Tages-Workshop) ▫ Nach-Trainingsphase (1 Jahr): <ul style="list-style-type: none"> - Implementierung von 4 durch den Auditausschuss organisierten Auditzyklen mit externer Unterstützung über jeweils 3 Monate ▫ Postinterventionsphase (1 Jahr): <ul style="list-style-type: none"> - Ermutigung des Gesundheitsfachpersonals zur Weiterführung der klinischen Audits ohne Supervision 	<p>gering</p>	<p>2</p>

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen
(Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungspotenzial auf Studienebene ^d	Evidenzniveau nach OCEBM
		<ul style="list-style-type: none"> - Zustand nach Kaiserschnitt: 11,4 % vs. 11,5 % - Risikoschwangerschaft: 52,9 % vs. 48,7 % ▫ 36 355 vs. 43 246 Gebärende in der Interventionsphase ▫ 23 484 vs. 28 781 Gebärende in der Postinterventionsphase 	<p>vs.</p> <ul style="list-style-type: none"> ▪ Keine Schulung des Personals (n = 16 Krankenhäuser, 100 725 Gebärende, davon 28 781 in der Postinterventionsphase) 		
<p>Lomas 1991 Cluster-RCT</p> <p>Ontario, Kanada 1988–1989</p> <p>Evaluation von Strategien zur Förderung der Umsetzung einer Leitlinie</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden <ul style="list-style-type: none"> ▫ Krankenhäuser ohne Status als Lehrinstitut mit mindestens 100 Betten, davon ≥ 10 geburtshilflich ▫ Gebärende mit 1 vorhergehenden niedrigen transversalen Kaiserschnitt, einer Einlingschwangerschaft in Schädellage und keiner absoluten Indikation für einen Kaiserschnitt (z. B. Plazenta praevia) 	<ul style="list-style-type: none"> ▪ 16 Krankenhäuser: 4 vs. 4 vs. 8 <ul style="list-style-type: none"> ▫ Geburtenzahl^e: 1230 ▫ Kaiserschnitt^e: 20,8 % ▫ Zustand nach Kaiserschnitt^e: 8,9 % ▫ VBAC^e: 3,1 % ▪ 2496 Gebärende <ul style="list-style-type: none"> ▫ Alter^e: 29,0 Jahre ▫ Parität^e: 1,15 	<ul style="list-style-type: none"> ▪ Schulung von Meinungsführern (n = 4 Krankenhäuser, 739 Gebärende) <ul style="list-style-type: none"> ▫ Wahl durch alle in der Geburtshilfe tätigen Ärztinnen und Ärzte einer Kollegin / eines Kollegen, die / der am besten zu den Beschreibungen eines einflussreichen Meinungsführers passt. Bereiterklärung der 4 benannten Ärzte, als Meinungsführer ihres Krankenhauses zu fungieren und an einem 1,5-tägigen Workshop zur Evidenz der Empfehlungen der Leitlinie und zu Grundprinzipien der Verhaltensänderung teilzunehmen. 	<p>hoch (wegen unklarer Angaben zur Erzeugung der Randomisierungssequenz und verdeckten Gruppenzuweisung, unklarer Verblindung der Gebärenden, unklarer ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

(Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungspotenzial auf Studienebene ^d	Evidenzniveau nach OCEBM
	<ul style="list-style-type: none"> ▪ Ausgeschlossen wurden <ul style="list-style-type: none"> ▫ Krankenhäuser, die in einem Landkreis mit einer Lehrinstitution oder im Norden Ontarios lagen ▫ Gebärende mit > 1 vorherigen Kaiserschnitt, einer vertikalen uterinen Narbe oder sonstigen Kontraindikationen, wie eine Beckenendlage, Zwillinge, eine Myomektomie, eine Abruption oder eine Plazenta praevia (entsprechend den Leitlinienkriterien) 		<ul style="list-style-type: none"> ▫ Versand einer Informationsmappe mit einem Begleitschreiben im Namen des Meinungsführers und detaillierten Informationen zur Leitlinie an das geburtshilflich tätige ärztliche Personal des jeweiligen Krankenhauses. ▫ 2-maliger Versand weiterer Detailinformationen, in denen Themen behandelt wurden, die von den Meinungsführern vereinbart wurden und für das Personal zur Umsetzung der Empfehlungen der Leitlinie von Interesse sein konnten ▫ Veranstaltung eines Treffens im Krankenhaus mit einem fachkundigen, kompetenten und glaubwürdigen Referenten im Bereich VBAC ▫ Aufrechterhaltung und Ausbau der regelmäßigen formellen und informellen Schulungskontakte mit Kolleginnen / Kollegen und Dokumentation dieser Kontakte für den Zeitraum der Intervention 		

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

(Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungspotenzial auf Studienebene ^d	Evidenzniveau nach OCEBM
			vs. <ul style="list-style-type: none"> ▪ Audit und Feedback (n = 4 Krankenhäuser, 524 Gebärende) <ul style="list-style-type: none"> ▫ Festlegung von Kriterien für die Anwendung des Kaiserschnitts bei Frauen mit einem vorhergehenden Kaiserschnitt, die auf der Leitlinie basieren, jedoch nicht unbedingt identisch mit dieser sein mussten ▫ ärztliche Überprüfung der Dokumentationsakten aller Frauen mit einem vorhergehenden Kaiserschnitt und Vergleich der tatsächlichen Praxis mit den vereinbarten Kriterien Besprechungen der Geburtshelferinnen und Geburtshelfer und in der Geburtshilfe tätigen Hausärztinnen und Hausärzte alle 3 Monate während des Jahres 1988 zur Diskussion der Audit-Ergebnisse und des Feedbacks, insbesondere der Diskrepanzen zwischen den festgelegten Kriterien und der tatsächlichen Praxis		

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

(Fortsetzung)

Quelle /Studien- typ / Rekrui- tierungsland / Rekrutierungs- zeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungs- potenzial auf Studienebene ^d	Evidenz- niveau nach OCEBM
			vs. <ul style="list-style-type: none"> ▪ Vergleichsintervention (n = 8 Krankenhäuser, 1233 Gebärende) ▫ Wiederholung des Verbreitungs- prozesses nach der ursprünglichen Veröffentlichung der Leitlinie Aktive Intervention nur zu Beginn der Studie, als allen in der Geburts- hilfe tätigen Personen eine Kopie der Leitlinie zugesandt wurde. Ein kurzes Anschreiben machte auf den Teil der Leitlinie aufmerksam, der sich mit der Anwendung des Kaisers- chnitts bei Frauen mit einem vor- herigen Kaiserschnitt befasste. Das Schreiben wies darauf hin, dass die Leitlinie von der nationalen Fach- gesellschaft für Geburtshilfe befür- wortet wurde, und forderte das geburtshilfliche Personal auf, die Empfehlungen umzusetzen. Darüber hinaus wurden keine weiteren Inter- ventionen vorgenommen. 		

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

(Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Intervention ^b (Prüfintervention) vs. Vergleichsintervention / Beginn / Umsetzung ^c	Verzerrungspotenzial auf Studienebene ^d	Evidenzniveau nach OCEBM
			<ul style="list-style-type: none"> ▪ Umsetzung: <ul style="list-style-type: none"> ▫ Schulung von Meinungsführern: <ul style="list-style-type: none"> - durchschnittliche Rücklaufquote für den Fragebogen zur Identifizierung des Meinungsführers 65,3 % - 3 Meinungsführer veranstalteten das Treffen mit einem Referenten. 1 Meinungsführer trat nach 6 Monaten zurück. - Die Aktivitäten der Meinungsführer waren bezogen auf Art und Umfang der Informationsvermittlung heterogen. ▫ Audit und Feedback: <ul style="list-style-type: none"> - In 2 der Krankenhäuser wurden alle erforderlichen Audit- und Feedback-Besprechungen durchgeführt. In den beiden anderen Krankenhäusern fanden 3 der 4 Treffen statt und das Personal erhielt per Post Feedback zum nicht stattgefundenen Treffen. - Feedback erfolgte ausschließlich auf Abteilungsebene ohne die Inanspruchnahme eines individuellen Feedbacks. 		

(Fortsetzung)

Tabelle 6: Charakteristika der eingeschlossenen Studie: Evidenztabelle zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

(Fortsetzung)

a: Extrahiert wurden alle berichteten Informationen zu Alter, Parität, Einlinge / Mehrlinge, SSW, Adipositas, vorheriger Kaiserschnitt, Informationen zur psychosozialen Belastung, Migrationshintergrund, Sozialstatus, PDA und Risikoprofil.

b: Die Angaben enthalten die Beschreibung der Weiterbildungsmaßnahmen in den Interventionsgruppen bzw. der Standardmaßnahme in der Vergleichsgruppe.

c: Einhaltung der Prüf- und Vergleichsintervention

d: ausführliche Beschreibung siehe Abschnitt 6.2

e: Mittelwert

n: Anzahl randomisierter Krankenhäuser bzw. Gebärender je Interventions- bzw. Vergleichsgruppe, für welche Ergebnisse vorliegen; PDA: Periduralanästhesie; OCEBM: Oxford Center for Evidence-Based Medicine; RCT: Randomized controlled Trial; SSW: Schwangerschaftswoche; VBAC: vaginale Geburt nach vorherigem Kaiserschnitt; vs.: versus

6.1.2 Übersicht der bewertungsrelevanten Endpunkte

Tabelle 7 zeigt die Übersicht der verfügbaren Daten zu den von der Leitliniengruppe festgelegten, für die Entscheidung kritischen 10 maternalen und 5 neonatalen Endpunkten aus der eingeschlossenen Studie. Die Endpunkte Spontangeburt, vaginal-operative Geburt und Kaiserschnitt wurden aufgrund der gegenseitigen Abhängigkeit voneinander zu dem Endpunkt Geburtsmodus zusammengefasst.

Zu den Endpunkten hoher mütterlicher Blutverlust ≥ 1000 ml, Stillbeziehung bei initialem Stillwunsch, teilweise / vollständige Plazentaretention, Geburtserleben, subjektive Gesundheitseinschätzung, Basendefizit ≥ -16 mmol/l und Verlegung auf die Neugeborenen-Intensivstation wurden keine Daten berichtet. Da der Apgar-Score nach 5 Minuten < 7 berichtet wurde, liegen zum kritischen Endpunkt Apgar-Score nach 5 Minuten < 6 keine Ergebnisse vor. Die vorliegenden ergänzenden Informationen zum Apgar-Score werden dargestellt, aber nicht in die Evidenzbewertung einbezogen.

Aus diesem Grund verbleiben Informationen zu 5 für die Entscheidung kritischen Endpunkten für die GRADE-Bewertung: maternale und neonatale Mortalität, Geburtsmodus, höhergradige Dammrisse (III° / IV°) sowie schwere metabolische Acidose. Zum Geburtsmodus liegen in 1 Studie ausschließlich Informationen zu der Anzahl der vaginalen Geburten vor. Eine Unterscheidung der vaginalen Geburten zwischen Spontangeburt und vaginal-operativen Geburten wurde in der Studie nicht vorgenommen.

Aus 1 der 2 eingeschlossenen Studien konnten Daten zu klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkten (Episiotomie, Wehenmittel, Amniotomie, Infektionen, fortgeschrittene Acidose) extrahiert werden. Zu den verbleibenden Endpunkten (Kristellerhilfe, Ikterus, Anämie) wurden in keiner Studie Daten berichtet.

Tabelle 7: Matrix der kritischen und klinisch wichtigen Endpunkte zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Kritische Endpunkte											Klinisch wichtige Endpunkte											
	Maternal											Neonatal				Maternal				Neonatal			
	Mortalität	Geburtsmodus			Höhergradige Dammrisse (III° / IV°)	Hoher mütterlicher Blutverlust ≥ 1000 ml	Stillbeziehung bei initialem Stillwunsch	Teilweise / vollständige Plazentaretention	Geburtserleben	Subjektive Gesundheitseinschätzung	Mortalität	Apgar-Score nach 5 Minuten < 6	Schwere metabolische Acidose	Basendefizit ≥ -16 mmol/l	Verlegung auf die NICU	Epistomie	Wehenmittel	Amniotomie	Kristellerhilfe	Infektionen	Leichte, mittelgradige, fortgeschrittene Acidose	Ikterus	Anämie
		Spontangeburt	Vaginal-operative Geburt	Kaiserschnitt																			
Chaillet 2015	■	○	■	● ^a	■	-	-	-	-	-	■	- ^b	■	-	-	■	■ ^c	■	-	■	■	-	-
Lomas 1991	■	■ ^d	■ ^e	-	-	-	-	-	-	■	-	-	-	-	-	-	-	-	-	-	-	-	

● primärer Endpunkt; ■ weitere / sekundäre Endpunkte; ○ aus berichteten Daten ableitbar; -: der Endpunkt wurde nicht erhoben
 a: Berichtet werden geplante Kaiserschnitte und ungeplante Kaiserschnitte im Geburtsverlauf.
 b: Es liegen ausschließlich Informationen zu einem Apgar-Score < 7 vor.
 c: Berichtet werden Informationen zur pharmakologischen Geburtseinleitung und Wehenmittelgabe unter der Geburt.
 d: Berichtet wird die Anzahl der vaginalen Geburten ohne Unterscheidung zwischen Spontangeburt und vaginal-operativer Geburt.
 e: Berichtet wird die Anzahl geplanter und ungeplanter Kaiserschnitte; es wurden ausschließlich Daten zu den ungeplanten Kaiserschnitten extrahiert.
 NICU: Neugeborenen-Intensivstation

6.2 Einschätzung des Verzerrungspotenzials

6.2.1 Einschätzung des Verzerrungspotenzials auf Studienebene

Für die Bewertung lagen 2 Studien vor. Die Einschätzung des Verzerrungspotenzials auf Studienebene ist in Tabelle 8 dargestellt.

Die Randomisierung erfolgte in beiden Studien clusterrandomisiert auf der Ebene der Krankenhäuser. In 1 Studie (Chaillet 2015) wurde die Randomisierung nach der Versorgungsstufe (Gemeinde-, Regional- oder Tertiärkrankenhaus) stratifiziert. Aufgrund der Angaben wird von einer adäquaten Erzeugung der Randomisierungssequenz und verdeckten Gruppenzuteilung ausgegangen. In der weiteren Studie (Lomas 1991) wurden in einem ersten Schritt 16 Verwaltungsbezirke in Nordontario zufällig ausgewählt und der Interventions- und Vergleichsgruppe zugeteilt. Aus diesen Verwaltungsbezirken wurden anschließend zufällig Krankenhäuser ausgewählt. Alle ausgewählten Krankenhäuser konnten eingeschlossen werden, sodass kein Rekrutierungsbias vorliegt. Es fehlen detailliertere Angaben zur Erzeugung der Randomisierungssequenz und der verdeckten Gruppenzuweisung. Deshalb wurde die adäquate Erzeugung der Randomisierungssequenz in dieser Studie mit unklar bewertet. Die Verdeckung der Gruppenzuteilung wird als unklar angesehen, da die Zuteilung der Krankenhäuser auf die Interventionsgruppen durch eine zentrale Einheit (die Studienkoordinatoren) erfolgte, deren Unabhängigkeit nicht beurteilt werden konnte.

Die Verblindung der teilnehmenden Gebärenden ist in beiden Studien unklar, da Angaben darüber fehlen, ob Gebärende über die zugeteilten Weiterbildungsmaßnahmen des geburtshilflichen Personals im jeweiligen Krankenhaus informiert waren. Aufgrund des Wissens, welcher Weiterbildungsmaßnahme das geburtshilfliche Personal im jeweiligen Krankenhaus zugeteilt war, lag bei den behandelnden Personen keine Verblindung vor. Diese fehlende Verblindung der behandelnden Personen wurde bei der Beurteilung des Verzerrungspotenzials auf Endpunktebene in Abschnitt 6.2.2 berücksichtigt.

In 1 Studie (Chaillet 2015) ist von einer ergebnisunabhängigen Berichterstattung auszugehen, während diese in der weiteren Studie (Lomas 1991) als unklar bewertet wurde, da für diese Studie keine Information zu den geplanten Endpunkten vorliegt und weder ein Studienprotokoll noch ein Studienregistereintrag verfügbar war. Es lagen keine sonstigen Aspekte vor.

In der Folge wurde das Verzerrungspotenzial auf Studienebene für 1 Studie mit gering und für 1 Studie mit hoch bewertet.

Tabelle 8: Bewertung des Verzerrungspotenzials auf Studienebene zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung der Teilnehmerinnen	Verblindung der behandelnden Personen	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
Chaillet 2015	ja	ja	unklar	nein	ja	ja	gering
Lomas 1991	unklar	unklar	unklar	nein	unklar	ja	hoch ^{a, b, c, d}

a: Angaben zur Erzeugung der Randomisierungssequenz fehlen.
b: Zuweisung erfolgte zentral durch die Studienkoordinatoren, deren Unabhängigkeit unklar ist.
c: Angaben, ob Gebärende über die zugeteilten Weiterbildungsmaßnahmen des ärztlichen Personals informiert waren, fehlen; dem teilnehmenden Personal der Krankenhäuser war die Gruppenzuteilung durch die Art der Intervention bekannt
d: unklare ergebnisunabhängige Berichterstattung, da kein Studienprotokoll vorliegt

6.2.2 Einschätzung des Verzerrungspotenzials auf Endpunktebene

Maternale Mortalität

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts maternale Mortalität wurde, bedingt durch das Verzerrungspotenzial auf Studienebene, in 1 Studie mit gering und in 1 Studie mit hoch bewertet. In 1 Studie (Chaillet 2015) basiert das Ergebnis auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 20 im Anhang).

Geburtsmodus

Das Verzerrungspotenzial für den Geburtsmodus basiert auf dem Verzerrungspotenzial auf Studienebene und endpunktspezifischen Abwertungen. Diese liegen in der fehlenden oder unklaren Verblindung der behandelnden und endpunkterhebenden Personen begründet, da diese einen Einfluss auf die Erfassung und Differenzierung eines geplanten oder ungeplanten Kaiserschnitts haben kann. In beiden Studien wurde die Clusterstruktur für 2 Ausprägungen (Lomas 1991) oder alle Ausprägungen (Chaillet 2015) nicht in der Ermittlung der Effektschätzung berücksichtigt. Zudem ist in 1 Studie (Lomas 1991) aufgrund eines intransparenten Patientenflusses die adäquate Umsetzung des ITT-Prinzips unklar, da Angaben zu absoluten Zahlen fehlen und widersprüchliche Angaben sowie Dateninkonsistenzen zu diesem Endpunkt vorliegen. Das Verzerrungspotenzial für den Geburtsmodus wurde folglich in 1 Studie mit mittel und in 1 Studie mit hoch bewertet (siehe Tabelle 21 im Anhang).

Höhergradige Dammrisse (III^o, IV^o)

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts höhergradige Dammrisse (III^o, IV^o) basiert auf dem Verzerrungspotenzial auf Studienebene und endpunktspezifischen Abwertungen. Diese liegen in der fehlenden oder unklaren Verblindung der behandelnden und endpunkterhebenden Personen begründet. Das Verzerrungspotenzial für höhergradige Dammrisse (III^o, IV^o) wurde folglich in Chaillet 2015 mit mittel bewertet. Das Ergebnis basiert auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt, bei denen davon auszugehen ist, dass kein höhergradiger Dammriss auftrat (siehe Tabelle 17 und Tabelle 22 im Anhang).

Neonatale Mortalität

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts neonatale Mortalität wurde, bedingt durch das Verzerrungspotenzial auf Studienebene, in Chaillet 2015 mit mittel und in Lomas 1991 mit hoch bewertet. In 1 Studie (Chaillet 2015) basiert das Ergebnis auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Aufgrund der wenigen Ereignisse konnte die Clusterstruktur im Modell in Chaillet 2015 nicht berücksichtigt werden. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 23 im Anhang).

Schwere metabolische Acidose

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts schwere metabolische Acidose wurde, bedingt durch das Verzerrungspotenzial auf Studienebene und fehlende Relevanz einer fehlenden Verblindung der behandelnden oder endpunkterhebenden Personen bei einer objektiven Messung dieses Endpunkts, in 1 Studie (Chaillet 2015) mit gering bewertet. Das Ergebnis basiert auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 24 im Anhang).

Episiotomie

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts Episiotomie basiert auf dem Verzerrungspotenzial auf Studienebene und endpunktspezifischen Abwertungen. Diese liegen in der fehlenden oder unklaren Verblindung der behandelnden und endpunkterhebenden Personen begründet. Das Verzerrungspotenzial für Episiotomie wurde folglich in 1 Studie (Chaillet 2015) mit mittel bewertet (siehe Tabelle 25 im Anhang).

Wehenmittel

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts Wehenmittel basiert auf dem Verzerrungspotenzial auf Studienebene und endpunktspezifischen Abwertungen. Diese liegen in der fehlenden oder unklaren Verblindung der behandelnden und endpunkterhebenden Personen begründet. Das Verzerrungspotenzial für Wehenmittelgaben unter der Geburt oder

zur Einleitung der Geburt wurde folglich in 1 Studie (Chaillet 20415) mit mittel bewertet (siehe Tabelle 26 und Tabelle 27 im Anhang).

Amniotomie

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts Amniotomie basiert auf dem Verzerrungspotenzial auf Studienebene und endpunktspezifischen Abwertungen. Diese liegen in der fehlenden oder unklaren Verblindung der behandelnden und endpunkterhebenden Personen begründet. Das Verzerrungspotenzial für Amniotomie wurde folglich in 1 Studie (Chaillet 2015) mit mittel bewertet. Das Ergebnis basiert auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 28 im Anhang).

Infektionen

Das Verzerrungspotenzial für die Effektschätzung des Endpunkts Infektionen wurde, bedingt durch das Verzerrungspotenzial auf Studienebene, in 1 Studie (Chaillet 2015) mit gering bewertet. Das Ergebnis basiert auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 29 im Anhang).

Leichte, mittelgradige und fortgeschrittene Acidose

In 1 Studie werden Ergebnisse für die fortgeschrittene Acidose berichtet. Das Verzerrungspotenzial für die Effektschätzung des Endpunkts fortgeschrittene Acidose wurde, bedingt durch das Verzerrungspotenzial auf Studienebene, in 1 Studie (Chaillet 2015) mit mittel bewertet. Aufgrund der wenigen Ereignisse konnte die Clusterstruktur im Modell nicht berücksichtigt werden. Das Ergebnis basiert auf allen Gebärenden, einschließlich der Frauen mit geplantem Kaiserschnitt. Die daraus resultierende mögliche Verzerrung wird in der Bewertung der Qualität der Evidenz unter Indirektheit berücksichtigt (siehe Tabelle 17 und Tabelle 30 im Anhang).

6.3 Ergebnisse zu kritischen Endpunkten

Die nachfolgende Tabelle 9 gibt einen Überblick über die Ergebnisse der Evidenz zu den kritischen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen.

Tabelle 9: Zusammenfassung der Ergebnisse zu den kritischen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Endpunkt	Ergebnisse
maternale Mortalität	aOR: 2,08; 95 %-KI: [0,05; 81,6]
Geburtsmodus	<ul style="list-style-type: none"> ▪ vaginale Geburt: OR: 1,07; 97,5 %-KI: [1,00; 1,14] ▪ Kaiserschnitt: OR: 0,89; 97,5 %-KI: [0,78; 1,01]
höhergradige Dammrisse (III°, IV°)	aOR: 1,07; 95 %-KI: [0,90; 1,26]
neonatale Mortalität	aOR: 0,34; 95 %-KI: [0,18; 0,65]
schwere metabolische Acidose (pH < 7,0)	aOR: 1,13; 95 %-KI: [0,79; 1,62]
aOR: adjustiertes Odds Ratio; KI: Konfidenzintervall; OR: Odds Ratio	

6.3.1 Maternale Mortalität

Zum Endpunkt maternale Mortalität lagen Daten aus 2 Studien vor.

In Chaillet 2015 verstarben insgesamt 4 (3,8 von 100 000) von 104 759 Gebärenden in der Prä- und Postinterventionsphase: 6,3 pro 100 000 in der Gruppe mit Prüfintervention verglichen mit 1,8 pro 100 000 in der Gruppe mit der Vergleichsintervention (OR: 2,08; 95 %-KI: [0,05; 81,6]). Dieses Ergebnis basiert auf den Daten aller 104 759 Gebärender in der Prä- und Postinterventionsphase, von denen insgesamt 13 419 (12,8 %) per geplantem Kaiserschnitt gebären. Informationen zu den Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

In Lomas 1991 traten bei insgesamt 504 Frauen in den Gruppen der Prüfintervention (Audit und Feedback bzw. Schulung von Meinungsführern) und bei 410 Frauen in der Gruppe der Vergleichsintervention kein Ereignis auf. Für dieses seltene Ereignis konnte weder für den Vergleich zwischen der Prüfintervention einer Schulung von Meinungsführern und der Vergleichsintervention (RD: 0,00; 95 %-KI: [-0,01; 0,01]) noch für den Vergleich zwischen der Prüfintervention mit Audit und Feedback und der Vergleichsintervention (RD: 0,00; 95 %-KI: [-0,01; 0,02]) ein statistisch signifikanter Zusammenhang nachgewiesen werden (siehe Tabelle 36 im Anhang).

Eine metaanalytische Zusammenfassung der Ergebnisse der 2 Studien erfolgte nicht, da das seltene Ereignis in Lomas 1991 aufgrund der kleinen Fallzahl von 2496 Frauen erwartungsgemäß nicht auftrat und ein Effekt aus dieser Studie nicht verlässlich geschätzt werden konnte. Deshalb erfolgt die Evidenzbewertung auf Grundlage des nicht signifikanten Ergebnisses aus der wesentlich größeren Studie Chaillet 2015 (OR: 2,08; 95 %-KI: [0,05; 81,6]).

6.3.2 Geburtsmodus

Zum Endpunkt Geburtsmodus lagen Daten aus 2 Studien vor. Diese berichten Ergebnisse zu 2 beziehungsweise 3 möglichen Ausprägungen des nominalen Endpunkts Geburtsmodus. Da

diese Ausprägungen wie separate binäre Endpunkte ausgewertet wurden, ergab sich aufgrund ihrer gegenseitigen Abhängigkeit voneinander die Notwendigkeit, das Signifikanzniveau der Analysen für multiples Testen zu adjustieren.

In Chaillet 2015 werden 3 mögliche Ausprägungen für den Geburtsmodus berichtet: Spontangeburt, vaginal-operative Geburt und Kaiserschnitt. Es wurde die Bonferroni-Methode verwendet und das 98,3 %-KI wurde berechnet.

In Lomas 1991 werden Ergebnisse aus 2 Vergleichen berichtet, wobei die Anzahl der vaginalen Geburten ohne Differenzierungsmöglichkeit zwischen Spontangeburt und vaginal-operativer Geburt berichtet wurde, sodass ausschließlich Ergebnisse zu den 2 Ausprägungen vaginale Geburten und Kaiserschnitt berichtet werden können. Zusätzlich wurde berücksichtigt, dass 2 Vergleiche zur Vergleichsintervention durchgeführt wurden. Es wurde die Bonferroni-Methode verwendet und das 98,75 %-KI für die Ereignisse vaginale Geburt und Kaiserschnitt beider Vergleiche berechnet.

Die Autorengruppen beider Studien berichten Ergebnisse zu der Anzahl ungeplanter und geplanter Kaiserschnitte. In Chaillet 2015 trat in der Postinterventionsperiode bei 12,2 % der Gebärenden der Gruppe mit der Prüfindervention ein geplanter Kaiserschnitt auf, während in der Gruppe mit der Vergleichsintervention bei 13,6 % der Gebärenden ein geplanter Kaiserschnitt auftrat. In Lomas 1991 trat ein geplanter Kaiserschnitt bei 53,7 %, 69,7 % und 66,8 % der Gebärenden in den Interventionsgruppen mit Schulungen von Meinungsführern, Audit und Feedback und der Gruppe mit der Vergleichsintervention auf. Für diesen Evidenzbericht wurde ausschließlich die Anzahl ungeplanter Kaiserschnitte extrahiert (siehe Tabelle 37 im Anhang).

Spontangeburt

Informationen zur Spontangeburt lagen aus 1 Studie (Chaillet 2015) vor. Insgesamt 35 542 (78,1 %) der 45 486 Gebärenden ohne geplanten Kaiserschnitt gebären in der Postinterventionsphase spontan. Es konnten keine signifikanten Unterschiede zwischen den Interventionsgruppen nachgewiesen werden: 78,3 % in der Gruppe mit Prüfindervention verglichen mit 78,0 % in der Gruppe mit der Vergleichsintervention gebären spontan (OR: 1,01; 98,3 %-KI: [0,96; 1,07]) (siehe Tabelle 37 im Anhang).

Vaginal-operative Geburt

Informationen zur vaginal-operativen Geburt lagen aus 1 Studie (Chaillet 2015) vor. Insgesamt 4828 (10,6 %) der 45 486 Gebärenden ohne geplanten Kaiserschnitt wiesen in der Postinterventionsphase eine vaginal-operative Geburt auf. Es konnten keine signifikanten Unterschiede zwischen den Interventionsgruppen nachgewiesen werden: 10,8 % in der Gruppe mit Prüfindervention verglichen mit 10,5 % in der Gruppe mit der Vergleichsintervention wiesen eine vaginal-operative Geburt auf (OR: 0,88; 98,3 %-KI: [0,75; 1,03]) (siehe Tabelle 37 im Anhang).

Vaginale Geburt

Informationen zur vaginalen Geburt lagen aus 2 Studien vor. Insgesamt 40 798 (87,9 %) der 46 400 Gebärenden ohne geplanten Kaiserschnitt gebären vaginal: 88,1 % in der Gruppe mit Prüfintervention verglichen mit 87,8 % in der Gruppe mit der Vergleichsintervention.

In Chaillet 2015 gebären in der Postinterventionsphase geringfügig mehr der 45 486 Gebärenden ohne geplanten Kaiserschnitt vaginal: 89,1 % der Gebärenden in der Gruppe mit Prüfintervention verglichen mit 88,5 % in der Gruppe mit der Vergleichsintervention (siehe Tabelle 37 im Anhang).

In Lomas 1991 gebären insgesamt 428 (46,8 %) der 914 Gebärenden mit vorherigem Kaiserschnitt, bei welchen laut Leitlinie ein vaginaler Geburtsversuch möglich war und durchgeführt wurde, vaginal: 54,2 % in der Interventionsgruppe mit Schulungen von Meinungsführern, 39,0 % der Gebärenden in der Interventionsgruppe mit Audit und Feedback verglichen mit 43,7 % in der Gruppe mit der Vergleichsintervention. Die Autorinnen und Autoren beschreiben einen signifikanten Unterschied zwischen den 3 Gruppen ($p = 0,003$). Für den Vergleich zwischen der Prüfintervention einer Schulung von Meinungsführern und der Vergleichsintervention konnte ein statistisch signifikanter Unterschied (RR: 1,24; 98,75 %-KI: [1,03; 1,50]) nachgewiesen werden, aber nicht zwischen der Prüfintervention mit Audit und Feedback und der Vergleichsintervention (RR: 0,89; 98,75 %-KI: [0,67; 1,19]) (siehe Tabelle 38 im Anhang).

Eine Metaanalyse mithilfe des Mantel-Haenszel-Modells konnte einen signifikanten Unterschied mit mehr vaginalen Geburten in der Gruppe mit der Prüfintervention (OR: 1,07; 97,5 %-KI: [1,00; 1,14]) bei nicht bedeutsamer Heterogenität der Ergebnisse der Einzelstudien nachweisen ($I^2 = 39,3 \%$, $p = 0,199$) (siehe Abbildung 2 im Anhang).

Kaiserschnitte

Informationen zum ungeplanten Kaiserschnitt lagen aus 2 Studien vor. Insgesamt 5602 (12,1 %) der 46 400 Gebärenden ohne geplanten Kaiserschnitt gebären per Kaiserschnitt im Geburtsverlauf: 11,9 % in der Gruppe mit der Prüfintervention verglichen mit 12,2 % in der Gruppe mit der Vergleichsintervention.

In Chaillet 2015 gebären in der Postinterventionsphase insgesamt 5116 (11,2 %) der 45 486 Gebärenden ohne geplanten Kaiserschnitt per Kaiserschnitt: 10,9 % der Gebärenden in der Gruppe mit Prüfintervention verglichen mit 11,5 % in der Gruppe mit der Vergleichsintervention (OR: 0,91; 98,3 %-KI: [0,78; 1,06]) (siehe Abbildung 3 im Anhang).

In Lomas 1991 gebären insgesamt 486 (53,2 %) der 914 Gebärenden mit vorherigem Kaiserschnitt, bei welchen laut Leitlinie ein vaginaler Geburtsversuch indiziert war, bei initial geplanter vaginaler Geburt per Kaiserschnitt im Geburtsverlauf: 45,8 % in der Interventionsgruppe mit Schulungen von Meinungsführern, 61,0 % der Gebärenden in der Interventionsgruppe mit Audit und Feedback verglichen mit 56,3 % in der Gruppe mit der

Vergleichsintervention. Die Autorinnen und Autoren konnten keinen signifikanten Unterschied zwischen den 3 Gruppen ($p = 0,166$) nachweisen. Für den Vergleich zwischen der Prüfindervention mit Schulungen von Meinungsführern und der Vergleichsintervention konnte ein statistisch signifikanter Unterschied (RR: 0,81; 98,75 %-KI: [0,67; 0,98]) nachgewiesen werden, aber nicht für den Vergleich zwischen der Prüfindervention mit Audit und Feedback und der Vergleichsintervention (RR: 1,08; 98,75 %-KI: [0,89; 1,31]) (siehe Tabelle 38 im Anhang).

Eine Metaanalyse mithilfe des inversen Varianz-Modells konnte bei nicht bedeutsamer Heterogenität der Ergebnisse der Einzelstudien ($I^2 = 0 \%$, $p = 0,341$) keinen signifikanten Unterschied zwischen den Interventionsgruppen nachweisen (OR: 0,89; 97,5 %-KI: [0,78; 1,01], siehe Abbildung 3 im Anhang).

6.3.3 Höhergradige Dammrise (III^o, IV^o)

Zum Endpunkt höhergradige Dammrise (III^o, IV^o) lagen Daten aus 1 Studie vor.

Insgesamt 2835 (6,2 %) der 45 486 Gebärenden ohne geplanten Kaiserschnitt erlitten in der Postinterventionsphase einen höhergradigen Dammriss: 6,0 % in der Gruppe mit der Prüfindervention verglichen mit 6,4 % in der Gruppe mit der Vergleichsintervention (OR: 1,07; 95 %-KI: [0,90; 1,26]) (siehe Tabelle 39 im Anhang).

6.3.4 Neonatale Mortalität

Zum Endpunkt neonatale Mortalität lagen Daten aus 2 Studien vor.

In Chaillot 2015 verstarben 48 (0,09 %) von 53 113 Neugeborenen in der Postinterventionsphase: 0,08 % in der Gruppe mit der Prüfindervention verglichen mit 0,1 % in der Gruppe mit der Vergleichsintervention (OR: 0,34; 95 %-KI: [0,18; 0,65]). Dieses Ergebnis basiert auf den Daten von insgesamt 52 265 Geburten, von denen insgesamt 6779 (13,0 %) per geplantem Kaiserschnitt erfolgten. Informationen zu den Neugeborenen von Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

In Lomas 1991 trat bei den Neugeborenen der insgesamt 504 Frauen in der Gruppe der Prüfindervention und bei den Neugeborenen der insgesamt 410 Frauen in der Gruppe der Vergleichsintervention kein Ereignis auf. Für dieses seltene Ereignis konnte weder für den Vergleich zwischen der Prüfindervention einer Schulung von Meinungsführern und der Vergleichsintervention (RD: 0,00; 95 %-KI: [-0,01; 0,01]) noch für den Vergleich zwischen der Prüfindervention mit Audit und Feedback und der Vergleichsintervention (RD: 0,00; 95 %-KI: [-0,01; 0,02]) ein statistisch signifikanter Zusammenhang nachgewiesen werden (siehe Tabelle 40 und Tabelle 41 im Anhang).

Eine metaanalytische Zusammenfassung erfolgte nicht, da dieses seltene Ereignis in Lomas 1991 aufgrund der kleinen Fallzahl bei 2496 Frauen erwartungsgemäß nicht auftrat und kein adjustiertes Ergebnis nachberechnet werden konnte. Deshalb wird das signifikante Ergebnis aus

Chaillet 2015 mit weniger neonatalen Todesfällen in der Gruppe mit der Prüfintervention interpretiert (OR: 0,34; 95 %-KI: [0,18; 0,65]).

6.3.5 Apgar-Score nach 5 Minuten < 6

Für den Apgar-Score nach 5 Minuten von < 6 liegen ausschließlich ergänzende Informationen vor (siehe Tabelle 42 im Anhang).

6.3.6 Schwere metabolische Acidose (pH < 7,0)

Zum Endpunkt schwere metabolische Acidose (pH < 7,0) lagen Daten aus 1 Studie vor. Insgesamt 177 (0,33 %) der 53 113 Neugeborenen wiesen in der Postinterventionsphase eine schwere metabolische Acidose auf: 0,36 % in der Gruppe mit der Prüfintervention verglichen mit 0,31 % in der Gruppe mit der Vergleichsintervention (OR: 1,13; 95 %-KI: [0,79; 1,62]) (siehe Tabelle 43 im Anhang). Dieses Ergebnis basiert auf den Daten von insgesamt 52 265 Geburten, von denen insgesamt 6779 (13,0 %) per geplantem Kaiserschnitt erfolgten. Informationen zu den Neugeborenen von Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

6.4 Ergebnisse zu wichtigen Endpunkten

Die nachfolgende Tabelle 10 gibt einen Überblick über die Ergebnisse der Evidenz zu den wichtigen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen für die eingeschlossenen Studien.

Tabelle 10: Zusammenfassung der Ergebnisse zu den wichtigen Endpunkten zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Endpunkt	Ergebnisse
Episiotomie	aOR: 1,01; 95 %-KI: [0,85; 1,21]
Wehenmittel	<ul style="list-style-type: none"> ▪ Wehenmittelgabe während der Geburt <ul style="list-style-type: none"> ▫ aOR: 1,16, 95 %-KI: [1,09; 1,23] ▪ Wehenmittelgabe zur Einleitung der Geburt <ul style="list-style-type: none"> ▫ aOR: 0,82, 95 %-KI: [0,76; 0,87]
Amniotomie	aOR: 1,02, 95 %-KI: [0,84; 1,22]
Infektionen	aOR: 0,95, 95 %-KI: [0,75; 1,19]
Fortgeschrittene metabolische Acidose	aOR: 1,49, 95 %-KI: [1,02; 2,17]
aOR: adjustiertes Odds Ratio; KI: Konfidenzintervall	

6.4.1 Episiotomie

Zum Endpunkt Episiotomie lagen Daten aus 1 Studie vor. Insgesamt 6824 (15,0 %) der 45 486 Gebärenden in der Postinterventionsphase, bei denen kein geplanter Kaiserschnitt vorgesehen war, erhielten eine Episiotomie: 14,3 % in der Gruppe mit der Prüfintervention

verglichen mit 15,6 % in der Gruppe mit der Vergleichsintervention (OR: 1,01; 95 %-KI: [0,85; 1,21]) (siehe Tabelle 44 im Anhang).

6.4.2 Wehenmittel

Zum Endpunkt Wehenmittel lagen Daten aus 1 Studie vor, in welcher Informationen zu Wehenmitteln während der Geburt und zur Einleitung der Geburt vorlagen. Insgesamt 13 777 (30,2 %) der 45 486 Gebärenden in der Postinterventionsphase, bei denen kein geplanter Kaiserschnitt vorgesehen war, erhielten eine Wehenmittelgabe während der Geburt: 30,1 % in der Gruppe mit der Prüfintervention verglichen mit 30,4 % in der Gruppe mit der Vergleichsintervention (OR: 1,16; 95 %-KI: [1,09; 1,23]) (siehe Tabelle 45 im Anhang). Insgesamt 13 373 (25,6 %) der 52 265 eingeschlossenen Gebärenden in der Postinterventionsphase erhielten eine Wehenmittelgabe zur Einleitung der Geburt: 23,4 % in der Interventionsgruppe verglichen mit 27,4 % in der Gruppe mit der Vergleichsintervention (OR: 0,82; 95 %-KI: [0,76; 0,87]) (siehe Tabelle 46 im Anhang).

6.4.3 Amniotomie

Zum Endpunkt Amniotomie lagen Daten aus 1 Studie vor. Insgesamt 26 506 (50,7 %) der 52 265 eingeschlossenen Gebärenden in der Postinterventionsphase erhielten eine Amniotomie: 51,0 % in der Gruppe mit der Prüfintervention verglichen mit 50,5 % in der Gruppe mit der Vergleichsintervention (OR: 1,02; 95 %-KI: [0,84; 1,22]) (siehe Tabelle 47 im Anhang). Dieses Ergebnis basiert auf den Daten aller 52 265 Gebärender, von denen insgesamt 6779 (13,0 %) per geplantem Kaiserschnitt gebären. Informationen zu den Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

6.4.4 Infektionen

Zum Endpunkt Infektionen lagen Daten aus 1 Studie vor. Insgesamt 4323 (8,3 %) der 52 265 eingeschlossenen Gebärenden in der Postinterventionsphase wiesen eine Infektion beziehungsweise Sepsis auf: 8,5 % in der Gruppe mit der Prüfintervention verglichen mit 8,1 % in der Gruppe mit der Vergleichsintervention (OR: 0,95; 95 %-KI: [0,75; 1,19]) (siehe Tabelle 48 im Anhang). Dieses Ergebnis basiert auf den Daten aller 52 265 Gebärender, von denen insgesamt 6779 (13,0 %) per geplanten Kaiserschnitt gebären. Informationen zu den Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

Zusätzlich lagen ergänzende Informationen zu neonatalen Infektionen vor (siehe Tabelle 49 im Anhang).

6.4.5 Fortgeschrittene metabolische Acidose

Zum Endpunkt fortgeschrittene metabolische Acidose (pH 7,09–7,0) lagen Daten aus 1 Studie vor. Insgesamt 809 (1,5 %) der 53 113 Neugeborenen wiesen in der Postinterventionsphase eine fortgeschrittene metabolische Acidose auf: 1,7 % in der Gruppe mit der Prüfintervention verglichen mit 1,4 % in der Gruppe mit der Vergleichsintervention (OR: 1,49; 95 %-KI: [1,02; 2,17]) (siehe Tabelle 50 im Anhang). Dieses Ergebnis basiert auf den Daten von insgesamt

52 265 Geburten, von denen insgesamt 6779 (13,0 %) per geplanten Kaiserschnitt erfolgten. Informationen zu den Neugeborenen von Gebärenden ohne geplanten Kaiserschnitt werden für diesen Endpunkt nicht berichtet.

7 Ergebnisse zur Erfahrung / Expertise durch systemimmanente Faktoren

7.1 Charakteristika der in die Bewertung eingeschlossenen Studien

7.1.1 Studiendesign und Studienpopulationen

Es wurden 3 Beobachtungsstudien auf der Basis von Daten der französischen nationalen Perinataldatenerhebungen (Belghiti 2013, Chuilon 2016, Coulm 2012) und 3 Beobachtungsstudien auf der Basis von nationalen Geburtsregistern (Finnström 2006, Karalis 2016 Wehberg 2018) zur Bewertung der Auswirkungen verschiedener Ausprägungen von Erfahrung und Expertise eingeschlossen, in denen Erfahrung / Expertise durch systemimmanente Faktoren beschrieben werden. Die Datenerfassungen wurden sowohl in den Perinataldatenerhebungen als auch in den Geburtsregistern prospektiv geplant und erfolgten nahezu vollständig, sodass von einer hohen externen Validität der Ergebnisse ausgegangen werden kann. Auf Grundlage der Beschreibungen in der Methodik der eingeschlossenen Studien wird von einer prospektiven Planung der Fragestellung und einer hypothesengeleiteten Auswertung in den eingeschlossenen Studien ausgegangen. Alle Ergebnisse basieren auf confounderadjustierten Auswertungen zu den unter 4.1.3 genannten Endpunkten.

Da keine randomisierten Studien identifiziert werden konnten, stammen alle berichteten Ergebnisse aus Kohortenstudien, in denen nicht verschiedene Interventionen, sondern Expositionen miteinander verglichen werden. Aus diesem Grund werden in der Untersuchung des Einflusses systemimmanenter Faktoren von Krankenhäusern keine Prüf- und Vergleichsinterventionen, sondern Expositionsgruppen (mit höheren Geburtenzahlen pro Jahr oder Versorgungsstufen der Geburtskliniken) mit Referenzgruppen (mit niedrigeren Geburtenzahlen / Versorgungsstufen der Geburtskliniken) verglichen.

Die Charakteristika der Studien auf der Grundlage der Perinataldatenerhebungen werden in Tabelle 11 und die Charakteristika der Studien auf der Grundlage der nationalen Geburtsregister werden in Tabelle 12 beschrieben.

Insgesamt wurden 2 079 155 Gebärende in die Studien eingeschlossen. Von diesen liegen für 1 891 204 Gebärende Ergebnisse vor. In 2 Studien (Chuilon 2016, Finnström 2006) wurden für insgesamt 1 550 136 Gebärende Ergebnisse für Frauen ohne Beschränkung auf Frauen mit einem geringen Risiko für Komplikationen berichtet, insgesamt 5,1 % der Frauen in Finnström 2006 gebären bereits vor der vollendeten 37. SSW. In Chuilon 2016 werden keine Informationen zum Gestationsalter berichtet.

Alle Studien untersuchten, ob eine größere Expertise im Vergleich zu einer geringeren Expertise, in der Regel beurteilt anhand der Geburtenzahlen pro Jahr der Geburtskliniken, einen Einfluss auf die Endpunkte von Mutter und Kind hat. Für keine dieser Studien liegen Informationen vor, ab welcher Phase der Geburt Informationen in den Perinatalerhebungen oder Geburtsregistern erhoben wurden.

3 Studien (Belghiti 2013, Chuilon 2016, Coulm 2012) basieren auf der französischen Perinataldatenerhebung im März 2010. In dieser Erhebung wurden alle Lebend- und Totgeburten in allen Geburtskliniken Frankreichs während einer vollen Woche im März 2010 erfasst, die nach der vollendeten 21. SSW oder mit einem Geburtsgewicht von ≥ 500 g geboren wurden. Dies betraf insgesamt 14 903 Geburten in insgesamt 535 Geburtskliniken. In den 3 Studien werden Ergebnisse für unterschiedliche Subgruppen berichtet, sodass jeweils eine unterschiedliche Anzahl an Gebärenden eingeschlossen wurde.

In Belghiti 2013 werden Ergebnisse für Frauen mit einem geringen Risiko für Interventionen während der Geburt mit einem lebenden Einling in Schädellage am errechneten Termin und einer für Frankreich beschriebenen Standardbetreuung berichtet. Die Standardbetreuung entspricht laut der Autorengruppe dem Betreuungsverlauf bei einer spontan eingesetzten Wehentätigkeit und einer PDA. Daten von 5668 Gebärenden wurden ausgewertet. Von diesen waren 43,0 % Erstgebärende und 57,0 % Mehrgebärende. Die Erfahrung / Expertise wird anhand unterschiedlicher Geburtenzahlen pro Jahr in den Geburtskliniken beschrieben. Größere Geburtskliniken mit ≥ 4000 Geburten pro Jahr wurden mit kleineren Geburtskliniken, eingeteilt in 4 Kategorien mit 3000–3999, 2000–2999, 1000–1999 beziehungsweise < 1000 Geburten pro Jahr, verglichen. Zudem wurden die Versorgungsstufen der Geburtskliniken verglichen, wobei spezialisierte Tertiärzentren mit einer Neugeborenen-Intensivstation (Level 3) die Exposition mit einer größeren Expertise darstellten und mit Geburtskliniken einer geringeren Versorgungsstufe (Krankenhäuser mit einer neonatalen Station mit Beatmungsmöglichkeit, einer neonatalen Station ohne Beatmungsmöglichkeit oder ohne eine Neugeborenenstation) verglichen wurden.

In Chuilon 2016 wurden ebenfalls Ergebnisse für alle Frauen nach vaginaler Geburt eines lebenden Kindes berichtet, im Gegensatz zu Belghiti 2013 wurde sich hier nicht auf die für Frankreich typische Standardbetreuung beschränkt. Es flossen Daten zu insgesamt 11 322 Frauen in diese Auswertung ein. Der Anteil an Erst- und Mehrgebärenden ist mit 42,9 % beziehungsweise 57,1 % nahezu identisch. Auch wurde die Erfahrung / Expertise anhand der unterschiedlichen Geburtenzahlen pro Jahr der jeweiligen Geburtskliniken sowie anhand der Versorgungsstufen der Geburtskliniken verglichen. Die Einteilung der Kategorien der Geburtskliniken war identisch zu Belghiti 2013.

In Coulm 2012 wurden ebenfalls Frauen mit einem geringen Risiko für Komplikationen eingeschlossen, die nach der 37 + 0. SSW einen lebenden Einling in Schädellage gebären. Insgesamt lagen Daten zu 9530 Gebärenden vor, wobei nicht für alle berichteten Endpunkte Informationen zu allen eingeschlossenen Frauen vorliegen. Auch hier wurde die Erfahrung / Expertise anhand der unterschiedlichen Geburtenzahlen pro Jahr der jeweiligen Geburtskliniken beschrieben, jedoch wurden die beiden höchsten Kategorien der Geburtskliniken (3000–3999 und ≥ 4000 Geburten pro Jahr) zusammengefasst. Die größere Expertise wurde in der Expositionsgruppe durch größere Geburtskliniken, eingeteilt in 3 Kategorien mit ≥ 3000 , 2000–2999 beziehungsweise 1000–1999 Geburten pro Jahr mit der Vergleichsgruppe, bestehend aus kleineren Geburtskliniken mit < 1000 Geburten pro Jahr, verglichen. Weiterhin

wurde ebenfalls die Expertise anhand der Versorgungsstufen der Geburtskliniken verglichen. Die Einteilung der Kategorien war hierbei identisch zu Belghiti 2013 und Chuilon 2016.

Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
<p>Belghiti 2013 Beobachtungsstudie auf der Basis der französischen Perinataldatenerhebung</p> <p>Frankreich März 2010</p> <p>Einfluss von individuellen Merkmalen der Frauen und Geburtskliniken auf die Häufigkeit einer Oxytocin-gabe während des Geburtsprozesses</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden Frauen mit einem geringen Risiko für Interventionen während der Geburt mit einem lebenden Einling in Schädellage am errechneten Termin und einer für Frankreich typischen Standardbetreuung (spontane Wehentätigkeit und PDA) ▪ Ausgeschlossen wurden Frauen mit <ul style="list-style-type: none"> ▫ chronischen Erkrankungen: <ul style="list-style-type: none"> - chronischem Bluthochdruck - Diabetes-Typ-1 - chronischem Organversagen - Thrombophilie - Lupus erythematoses - Antiphospholipid-Syndrom - Epilepsie - chronischen Infektionen mit Hepatitis B / C / HIV ▫ einer besonderen geburtshilflichen Anamnese: <ul style="list-style-type: none"> - vorheriger Kaiserschnitt - Schwangerschaftscholestase - Präeklampsie oder Totgeburt 	<ul style="list-style-type: none"> ▪ N = 5668 ▪ Parität: <ul style="list-style-type: none"> ▫ Erstgebärende: 43,0 %^d ▫ Mehrgebärende: 57,0 %^d 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand der unterschiedlichen Geburtzahlen der Krankenhäuser (n = 5668) <ul style="list-style-type: none"> ▫ größere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - ≥ 4000 (n = 256) vs. ▫ mittlere und kleinere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - 3000–3999 (n = 797) - 2000–2999 (n = 1745) - 1000–1999 (n = 1956) - < 1000 (n = 914) ▪ Beurteilung der Expertise anhand der Versorgungsstufe der Geburtsklinik (n = 5660): <ul style="list-style-type: none"> - Level 3: spezialisierte Tertiärzentren mit Neugeborenen-Intensivstation (n = 1261) vs. - Level 2b: Geburtsklinik mit einer neonatalen Station mit Beatmungsmöglichkeit (n = 1175) - Level 2a: Geburtsklinik mit einer neonatalen Station ohne Beatmungsmöglichkeit (n = 1565) 	<p>hoch (wegen Risiko für Selektionsbias aufgrund fehlender Randomisierung, fehlender Informationen zu Baseline-Charakteristika der Vergleichsgruppen, fehlender Verblindung und unklarer ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
	<ul style="list-style-type: none"> ▫ Erkrankungen in der aktuellen Schwangerschaft <ul style="list-style-type: none"> - schwere Präeklampsie - insulinpflichtiger Gestationsdiabetes - Rhesus-Inkompatibilität ▫ intrauterine Wachstumsretardierung < 5. Perzentil 		<ul style="list-style-type: none"> - Level 1: Geburtsklinik ohne Neugeborenen-Intensivstation (n = 1659) 		
<p>Chuilon 2016 Beobachtungsstudie auf der Basis der französischen Perinataldatenerhebung</p> <p>Frankreich März 2010</p> <p>Einfluss von individuellen Merkmalen von Frauen / Kindern und Geburtskliniken auf die Häufigkeit von Episiotomien</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden Erst- und Mehrgebärende ab der 22. SSW oder mit einem Neugeborenen mit einem Geburtsgewicht > 500 g. ▪ Ausgeschlossen wurden Frauen mit einer Entbindung per Kaiserschnitt, einer Mehrlingschwangerschaft, einem intrauterin verstorbenen Kind, einer Abruption oder bei fehlenden Informationen zur Episiotomie. 	<ul style="list-style-type: none"> ▪ N = 11 322 ▪ Parität^e: <ul style="list-style-type: none"> ▫ Erstgebärende: 42,9 %^d ▫ Mehrgebärende: 57,1 %^d ▪ Alter^e: <ul style="list-style-type: none"> ▫ < 20: 2,6 %^d ▫ 20–24: 15,2 %^d ▫ 25–29: 34,0 %^d ▫ 30–34: 30,7 %^d ▫ 35–39: 14,5 %^d ▫ > 40: 3,1 %^d ▪ BMI^e: <ul style="list-style-type: none"> ▫ < 18,5: 8,9 %^d ▫ 18,5–24,9: 66,0 %^d ▫ 25–29,9: 16,7 %^d 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand der unterschiedlichen Geburtenzahlen der Krankenhäuser (n = 11 142) <ul style="list-style-type: none"> ▫ größere Geburtsstationen gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - ≥ 4000 (n = 485) vs. ▫ mittlere und kleinere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - 3000–3999 (n = 1510) - 2000–2999 (n = 3248) - 1000–1999 (n = 3885) - < 1000 (n = 2014) ▪ Beurteilung der Expertise anhand der Versorgungsstufe der Geburtsklinik (n = 11 133): 	<p>hoch (wegen Risiko für Selektionsbias aufgrund fehlender Randomisierung, fehlender Informationen zu Baseline-Charakteristika der Vergleichsgruppen, fehlender Verblindung und unklarer ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
		<ul style="list-style-type: none"> ▫ ≥ 30: 8,5 %^d ▪ vorheriger Kaiserschnitt^e: 4,8 %^d ▪ Migrationshintergrund^{e, f} <ul style="list-style-type: none"> ▫ Frankreich: 82,7 %^d ▫ andere europäische Länder: 4,0 %^d ▫ Nordafrika: 6,9 %^d ▫ andere afrikanische Länder: 3,5 %^d ▫ übrige Länder: 3 %^d 	<ul style="list-style-type: none"> ▫ Level 3: spezialisiertes Tertiärzentren mit einer Neugeborenen-Intensivstation (n = 2463) vs. ▫ Level 2b: neonatale Station mit Beatmungsmöglichkeit (n = 2329) ▫ Level 2a: neonatale Station ohne Beatmungsmöglichkeit (n = 2983) ▫ Level 1: ohne Neugeborenen-Intensivstation (n = 3358) 		

(Fortsetzung)

Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
<p>Coulm 2012 Beobachtungsstudie auf der Basis der französischen Perinataldatenerhebung</p> <p>Frankreich März 2010</p> <p>Einfluss von Größe, Versorgungsstufe und Krankenträger einer Geburtsklinik auf das medizinische Handeln während des Geburtsprozesses</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden Erst- und Mehrgebärende mit einem geringen Risiko und einer termingerechten Geburt $\geq 37 + 0$ SSW eines lebenden Einlings in Schädellage. ▪ Ausgeschlossen wurden alle Frauen mit <ul style="list-style-type: none"> ▫ einem Risikofaktor wie einer schweren maternalen Erkrankung: <ul style="list-style-type: none"> - chronischem Bluthochdruck - Diabetes Typ 1 - chronischem Organversagen - Thrombophilie - Lupus erythematoses - Antiphospholipid-Syndrom - schwerer Epilepsie - seropositiv für Hepatitis B / C / HIV ▫ einer ungünstigen geburtshilflichen Anamnese <ul style="list-style-type: none"> - vorheriger Kaiserschnitt - Schwangerschaftscholestase - Präeklampsie oder 	<ul style="list-style-type: none"> ▪ N = 9530 ▪ Parität^e: <ul style="list-style-type: none"> ▫ Erstgebärende: 41,7 %^d ▫ Mehrgebärende: 58,3 %^d ▪ SSW^e: <ul style="list-style-type: none"> ▫ 37–38: 20,9 %^d ▫ 39–40: 57,4 %^d ▫ ≥ 41: 21,6 %^d ▪ Alter^e: <ul style="list-style-type: none"> ▫ < 25: 17,9 %^d ▫ 25–29: 34,6 %^d ▫ 30–34: 30,5 %^d ▫ ≥ 35: 17,0 %^d ▪ BMI^e: <ul style="list-style-type: none"> ▫ < 18,5: 8,4 %^d ▫ 18,5–24,9: 66,3 %^d ▫ 25–29,9: 16,7 %^d ▫ ≥ 30: 8,6 %^d ▪ Bildungsgrad^e: <ul style="list-style-type: none"> ▫ < Sekundarstufe I: 26,4 %^d ▫ Sekundarstufe II: 19,8 %^d 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand der unterschiedlichen Geburtzahlen der Krankenhäuser (n = 9523) <ul style="list-style-type: none"> ▫ größere und mittlere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - ≥ 3000 (n = 1671) - 2000–2999 (n = 2721) - 1000–1999 (n = 3370) vs. ▫ kleinere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - < 1000 (n = 1761) ▪ Beurteilung der Expertise anhand der Versorgungsstufe der Geburtsklinik (n = 9429) <ul style="list-style-type: none"> ▫ Geburtsklinik mit angeschlossener Kinderklinik <ul style="list-style-type: none"> - Level 3: spezialisierte Tertiärzentren mit einer Neugeborenen-Intensivstation (n = 2003) - Level 2b: neonatale Station mit Beatmungsmöglichkeit (n = 1915) - Level 2a: neonatale Station ohne Beatmungsmöglichkeit (n = 2499) vs. 	<p>hoch (wegen Risiko für Selektionsbias aufgrund fehlender Randomisierung, fehlender Informationen zu Baseline-Charakteristika der Vergleichsgruppen, fehlender Verblindung und unklarer ergebnisunabhängiger Berichterstattung)</p>	

(Fortsetzung)

Tabelle 11: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von nationalen Perinataldatenerhebungen (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
	<ul style="list-style-type: none"> - Totgeburt ▫ Erkrankungen in der aktuellen Schwangerschaft <ul style="list-style-type: none"> - schwere Präeklampsie - insulinpflichtiger Gestationsdiabetes - Rhesus-Inkompatibilität - intrauterine Wachstumsretardierung < 5. Perzentil 	<ul style="list-style-type: none"> ▫ akademische Bildung: 53,8, %^d ▪ Migrationshintergrund^{e, f}: <ul style="list-style-type: none"> ▫ Frankreich: 82,6 %^d ▫ andere europ. Länder: 3,8 %^d ▫ Nordafrika: 7,0 %^d ▫ andere afrikan. Länder: 3,2 %^d ▪ übrige Länder: 3,3 %^d 	<ul style="list-style-type: none"> ▫ Geburtsklinik ohne angeschlossene Kinderklinik <ul style="list-style-type: none"> - Level 1: ohne Neugeborenen-Intensivstation (n = 3012) 		
<p>a: Extrahiert wurden alle berichteten Informationen zu Alter, Parität, Einlinge / Mehrlinge, SSW, Adipositas, vorheriger Kaiserschnitt, Informationen zur psychosozialen Belastung, Migrationshintergrund, Sozialstatus, PDA und Risikoprofil.</p> <p>b: Die Angaben enthalten die Beschreibung der Beurteilung der größeren Erfahrung / Expertise in der Expositionsgruppe im Vergleich zur geringeren Erfahrung / Expertise in der Referenzgruppe, es können 1 oder mehrere Expositions- und Referenzgruppen auftreten.</p> <p>c: ausführliche Beschreibung siehe Abschnitt 5.3</p> <p>d: eigene Berechnung; Anzahl der berichteten Informationen der Gesamtangaben variiert</p> <p>e: Gesamtanzahl der Gebärenden mit vorliegenden Werten variiert zwischen den Charakteristika und weicht von der Gesamtanzahl ab.</p> <p>f: Geburtsland der Gebärenden</p> <p>BMI: Body-Mass-Index; N: Anzahl Gebärender; n: Anzahl Gebärender je Expositions- bzw. Referenzgruppe; PDA: Periduralanästhesie; SSW: Schwangerschaftswoche; vs.: versus</p>					

Insgesamt 3 Beobachtungsstudien basieren auf nationalen medizinischen Geburtsregistern aus Schweden (Finnström 2006), Finnland (Karalis 2016) und Dänemark (Wehberg 2018) und umfassen verschiedene Perioden im Zeitraum von 1985 bis 2012.

Finnström 2006 wertet die Daten von 1 538 814 lebenden Einlingsgeburten der Jahre 1985 bis 1999 aus dem schwedischen medizinischen Geburtsregister aus. Laut Angaben der Autorinnen und Autoren sind Daten zu 99 % aller Geburten in Schweden in diesem Register enthalten. 42,0 % der Gebärenden waren Erstgebärende und folglich 58,0 % Mehrgebärende. 94,7 % der Geburten fanden nach der vollendeten 37. SSW statt. Die Erfahrung / Expertise wird in Finnström 2006 anhand der Geburtenzahlen pro Geburtsklinik beurteilt, es werden Geburtskliniken verschiedener Größe (≥ 2500 , 500–999, < 500 Geburten pro Jahr) mit der Referenzkategorie von 1000–2499 Geburten pro Jahr verglichen. Die meisten Frauen entbinden in Geburtskliniken dieser Größe, welche mehrheitlich eine Kinderklinik enthalten. Geburtskliniken mit weniger als 500 Geburten beinhalten in der Regel keine angeschlossene Kinderklinik. Einige Geburtskliniken wechselten während des Untersuchungszeitraums die Gruppenzugehörigkeit, d. h. von mehr als 1000 Geburten pro Jahr auf unter 1000 Geburten pro Jahr. Dies wurde laut den Autorinnen und Autoren in den Analysen berücksichtigt. In Finnström 2006 werden zudem Ergebnisse nach dem Einzugsgebiet der Gebärenden berichtet. Das Einzugsgebiet wurde durch das Krankenhaus definiert, in dem mehr als 70 % der Frauen in einer bestimmten Gemeinde entbunden haben. Wenn mehr als 70 % der Frauen in verschiedenen Krankenhäusern, aber mit der gleichen Krankenhausgröße entbunden haben, wurden sie in das Einzugsgebiet dieser Größe eingeteilt. Die Ergebnisse werden in Sensitivitätsanalysen zur Abschätzung von Selektionsbias durch die Wahl des Krankenhauses oder die Überweisung von Frauen mit Risiken oder speziellem Behandlungsbedarf genutzt. Die Mehrheit der Gebärenden entschied sich laut Angaben der Autorengruppe für die Entbindung in ihrem Einzugsgebiet. Bei auftretenden Komplikationen während der Schwangerschaft war eine Überweisung der Schwangeren in eine Geburtsklinik mit einer höheren Versorgungsstufe vorgesehen.

Die Studie von Karalis 2016 basiert auf dem finnischen medizinischen Geburtsregister der Jahre 2005 bis 2009. Die Studienpopulation umfasst 267 390 Geburten, was laut den Angaben der Autorinnen und Autoren 90,2 % der Geburten in Finnland entspricht. Aufgrund fehlender Daten konnten 267 066 Geburten mit einem niedrigen Risiko ausgewertet werden. 41,7 % der Gebärenden waren Erstgebärende und folglich 58,2 % Mehrgebärende. Angaben zum Gestationsalter liegen in der Publikation nicht vor, Geburten mit einem Geburtsgewicht < 2500 g wurden aus der Analyse ausgeschlossen. Ausgehend von einer normalen Schwangerschaftsdauer von 40 Wochen (40 + 0 SSW) und einer normalverteilten Abweichung vom Geburtstermin, nach der nur 17,5 % aller Frauen mehr als 10 Tage vor dem errechneten Geburtstermin gebären [8], wurde davon ausgegangen, dass mindestens 80 % der Frauen nach der vollendeten 37. SSW gebären. Die Erfahrung / Expertise wurde auf der Basis des Versorgungsstatus der Geburtsklinik und der jährlichen Anzahl der Geburten der Geburtskliniken verglichen. Daher wurden Ergebnisse aus Universitätskrankenhäusern mit

denen von großen (≥ 2000 Geburten pro Jahr), mittleren (1000–1999 Geburten pro Jahr) und kleinen (≤ 999 Geburten pro Jahr) Geburtskliniken verglichen.

In Karalis 2016 liegen keine Angaben zum Gestationsalter vor. Ausgehend von einer normalen Schwangerschaftsdauer von 40 Wochen (40 + 0 SSW) und einer normalverteilten Abweichung vom Geburtstermin, nach der nur 17,5 % aller Frauen mehr als 10 Tage vor dem errechneten Geburtstermin gebären [8], wurde auf Grundlage der in der Studie genannten Einschlusskriterien davon ausgegangen, dass mindestens 80 % der Geburten nach der vollendeten 37. SSW stattfanden.

Die Studie von Wehberg 2018 basiert auf den Daten des dänischen medizinischen Geburtsregisters der Jahre 2009 bis 2012. In diesem Zeitraum wurden Daten für 229 041 Geburten berichtet. Insgesamt 226 612 Einlingsgeburten von insgesamt 198 590 Erstgebärenden mit einem geringen Risiko für Komplikationen, die nach der vollendeten 37. SSW gebären, wurden in die Studie eingeschlossen. Insgesamt 96,4 % der eingeschlossenen Gebärenden gebären zwischen der vollendeten 37. und 41. SSW. Die Erfahrung / Expertise wurde anhand der Geburtenzahlen pro Geburtsklinik beurteilt, sodass größere (> 3000 Geburten pro Jahr) mit mittleren Geburtskliniken (1100 bis 3000 Geburten pro Jahr) und kleineren Geburtskliniken mit < 1100 Geburten pro Jahr verglichen wurden.

Tabelle 12: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von Geburtsregistern

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
<p>Finnström 2006 Beobachtungsstudie auf der Basis des schwedischen medizinischen Geburtsregisters</p> <p>Schweden 1985–1999</p> <p>Einfluss der Größe der Geburtsklinik auf neonatale Endpunkte</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden lebende Einlingsgeburten. ▪ keine Angaben zu Ausschlusskriterien 	<ul style="list-style-type: none"> ▪ N = 1 538 814 ▪ Parität: <ul style="list-style-type: none"> ▫ Erstgebärende: 42,0 %^d ▫ Mehrgebärende: 58,0 %^d ▪ SSW: <ul style="list-style-type: none"> ▫ ≤ 32: 0,9 %^d ▫ 33–36: 4,2 %^d ▫ ≥ 37: 94,7 %^d ▫ k. A: 0,2 %^d ▪ Alter: <ul style="list-style-type: none"> ▫ < 20: 2,6 %^d ▫ 20–34: 84,9 %^d ▫ > 34: 12,5 %^d ▫ k. A: 906 (0,06 %)^d ▪ BMI: <ul style="list-style-type: none"> ▫ < 25: 49,8 %^d ▫ ≥ 25: 15,9 %^d ▫ k. A: 34,3 %^d 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand der unterschiedlichen Geburtenzahlen der Krankenhäuser <ul style="list-style-type: none"> ▫ größere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - ≥ 2500 (n = 239 441) vs. ▫ mittlere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - ≥ 1000–2499 (n = 638 183) vs. ▫ kleinere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: <ul style="list-style-type: none"> - 500–999 (n = 175 388) - < 500 (n = 31 745) ▪ Die meisten Krankenhäuser mit < 500 Geburten haben keine Kinderklinik, während die Mehrheit der Krankenhäuser > 1000 Geburten mit einer Kinderklinik ausgestattet sind. 	<p>hoch (wegen Risiko eines Selektionsbias, fehlender Verblindung, fehlender Informationen zu Baseline-Charakteristika, fehlender ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 12: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von Geburtsregistern (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
<p>Karalis 2016 Beobachtungsstudie auf der Basis des finnischen Nationalen Medizinischen Geburtsregisters</p> <p>Finland 2005–2009</p> <p>Einfluss der Größe der Geburtskliniken und der ärztlichen Rufbereitschaft auf neonatale Endpunkte</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden Geburten mit einem geringen Risiko. ▪ Ausgeschlossen wurden Geburten mit einem vorgeburtlich verstorbenen Fetus, geplante oder ungeplante Hausgeburten, Entbindungen auf dem Weg ins Krankenhaus, Mehrlingsschwangerschaften, Geburten mit einem Geburtsgewicht < 2500 g oder mit schweren kongenitalen Anomalien des Neugeborenen. 	<ul style="list-style-type: none"> ▪ Geburten insgesamt N = 296 397; Einschluss von N = 267 390 Geburten mit geringem Risiko; Auswertung von N = 267 066 Geburten mit vollständigen Daten ▪ Parität: <ul style="list-style-type: none"> ▫ Erstgebärende: 41,7 % ▫ Mehrgebärende: 58,2 %^d ▪ Alter: <ul style="list-style-type: none"> ▫ 14–19: 2,5 % ▫ 20–24: 16,2 % ▫ 25–29: 32,2 % ▫ 30–34: 31,0 % ▫ 35–39: 14,6 % ▫ 40–44: 3,5 %^d ▫ 45–52: 0,1 % 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand des Versorgungsstatus und der unterschiedlichen Geburtenzahlen der Krankenhäuser <ul style="list-style-type: none"> ▫ Universitätskrankenhaus (n = 92 139) vs. ▫ nicht Universitätskrankenhaus <ul style="list-style-type: none"> - größere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: ≥ 2000 (n = 84 681) - mittlere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: 1000–1999 (n = 51 185) - kleinere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: ≤ 999 (n = 39 385) 	<p>hoch (wegen Risiko eines Selektionsbias, fehlender Verblindung, fehlender Informationen zu Baseline-Charakteristika, fehlender ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 12: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von Geburtsregistern (Fortsetzung)

Quelle / Studientyp / Rekrutierungsland / Rekrutierungszeitraum / Ziel	Ein- / Ausschlusskriterien	Charakteristika ^a der Population	Exposition ^b (Expositionsgruppen) vs. Referenzgruppe	Verzerrungspotenzial auf Studienebene ^c	Evidenzniveau nach OCEBM
<p>Wehberg 2018 Beobachtungsstudie auf der Basis des Dänischen Medizinischen Geburtsregisters</p> <p>Dänemark 2009–2012</p> <p>Einfluss von verschiedenen Risikofaktoren auf einen geplanten oder ungeplanten Kaiserschnitt</p>	<ul style="list-style-type: none"> ▪ Eingeschlossen wurden lebende und totegeborene Kinder mit einem Gestationsalter ab 24 +0 SSW. Berichtet werden die Ergebnisse für Erstgebärende ab der vollendeten 37. SSW mit einem geringen Risiko, die ein Kind in Schädellage mit einer spontanen Wehentätigkeit gebären. ▪ Ausgeschlossen wurden Geburten in sehr kleinen Abteilungen mit < 100 Geburten pro Jahr, ungültig dokumentierter Kennzahl der Geburtsklinik, einer fehlenden / inkorrekten Geburtsdiagnose bzw. einem unschlüssigen Geburtsmodus, nach einem induzierten Abort und Frauen jünger als 15 oder älter als 44 Jahre. 	<ul style="list-style-type: none"> ▪ Geburten insgesamt: N = 229 041 Einschluss von N = 226 612 Einlingsgeburten von 198 590 Erstgebärenden ▪ Auswertung einer Subgruppe von Erstgebärenden mit einem geringen Risiko: N = 74 002 ▪ SSW^f: <ul style="list-style-type: none"> ▫ 37 + 0–41 + 6: 96,4 %^d ▫ ≥ 42 + 0: 3,6 %^d ▪ Alter^f <ul style="list-style-type: none"> ▫ 15–19: 3,1 %^d ▫ 20–34: 86,6 %^d ▫ 35–44: 10,3 %^d ▪ BMI^f: <ul style="list-style-type: none"> ▫ 14,0–18,4: 5,0 %^d ▫ 18,5–24,9: 68,1 %^d ▫ 25,0–29,9: 18,2 %^d ▫ 30–34,9: 6,2 %^d ▫ ≥ 35,0: 2,7 %^d ▪ Komplikationen vorhanden^{f, g}: 10,5 %^e 	<ul style="list-style-type: none"> ▪ Beurteilung der Expertise anhand der unterschiedlichen Geburtenzahlen der Krankenhäuser <ul style="list-style-type: none"> ▫ größere Geburtsstationen gemessen an der Anzahl der Geburten pro Jahr: - ≥ 3000 (n = 41 689) vs. ▫ mittlere Geburtskliniken gemessen an der Anzahl der Geburten pro Jahr: - 1100–3000 (n = 27 970) vs. ▫ kleinere Geburtsklinik gemessen an der Anzahl der Geburten pro Jahr: - < 1100 (n = 4343) 	<p>hoch (wegen Risiko eines Selektionsbias, fehlender Verblindung, fehlender Informationen zu Baseline-Charakteristika, unklarer ergebnisunabhängiger Berichterstattung)</p>	<p>3</p>

(Fortsetzung)

Tabelle 12: Charakteristika der eingeschlossenen Studien: Evidenztabelle zur Erfahrung / Expertise durch systemimmanente Faktoren auf der Grundlage von Geburtsregistern (Fortsetzung)

a: Extrahiert wurden alle berichteten Informationen zu Alter, Parität, Einlinge / Mehrlinge, SSW, Adipositas, vorheriger Kaiserschnitt, Informationen zur psychosozialen Belastung, Migrationshintergrund, Sozialstatus, PDA und Risikoprofil.
b: Die Angaben enthalten die Beschreibung der Beurteilung der größeren Erfahrung / Expertise in der Expositionsgruppe im Vergleich zur geringeren Erfahrung / Expertise in der Referenzgruppe, es können 1 oder mehrere Expositions- und Referenzgruppen auftreten.
c: ausführliche Beschreibung siehe Abschnitt 7.2.1
d: eigene Berechnung
e: Originalangabe aus der Publikation, gemeint ist wahrscheinlich ≥ 5000
f: Berichtete Daten beziehen sich auf Subgruppe der Erstgebärende um den errechneten Termin.
g: Es liegen keine weiteren Angaben zur Art der Komplikation vor.
BMI: Body-Mass-Index; N: Anzahl Gebärender; n: Anzahl Gebärender je Expositions- bzw. Referenzgruppe; PDA: Periduralanästhesie; SSW: Schwangerschaftswoche; vs.: versus

7.1.2 Übersicht der bewertungsrelevanten Endpunkte

Tabelle 13 zeigt die Übersicht der verfügbaren Daten zu den von der Leitliniengruppe festgelegten, für die Entscheidung kritischen 10 maternalen und 5 neonatalen Endpunkten aus der eingeschlossenen Studie. Die Endpunkte Spontangeburt, vaginal-operative Geburt und Kaiserschnitt wurden aufgrund der gegenseitigen Abhängigkeit voneinander zu dem Endpunkt Geburtsmodus zusammengefasst.

Zu den kritischen Endpunkten maternale Mortalität, höhergradige Dammmrisse (III° / IV°), hoher mütterlicher Blutverlust ≥ 1000 ml, Stillbeziehung bei initialem Stillwunsch, teilweise / vollständige Plazentaretention, Geburtserleben, subjektive Gesundheitseinschätzung und Basendefizit ≥ -16 mmol/l wurden keine Daten berichtet. Da der Schwellenwert für einen auffälligen Apgar-Score nach 5 Minuten in 2 Studien bei < 4 und in 1 Studie < 7 angesetzt wurde, liegen zum kritischen Endpunkt Apgar-Score nach 5 Minuten < 6 keine Ergebnisse vor. In 2 Studien werden Informationen zu Neugeborenen mit einer Verlegung auf eine Neugeborenen-Intensivstation, eine Beobachtungsstation / Special Care Nursery oder in ein anderes Krankenhaus berichtet, ohne dass zwischen einer Verlegung auf die Neugeborenen-Intensivstation und andere Stationen unterschieden werden kann. Die vorliegenden ergänzenden Informationen zum Apgar-Score und zur Verlegung des Neugeborenen werden dargestellt, aber nicht in die Evidenzbewertung einbezogen.

Es verbleiben Informationen zu 3 für die Entscheidung kritischen Endpunkten für die GRADE-Bewertung: neonatale Mortalität, Geburtsmodus und schwere metabolische Acidose. Zum Geburtsmodus liegen in 1 Studie ausschließlich Informationen zu der Gesamtanzahl der Kaiserschnitte ohne eine Differenzierungsmöglichkeit zwischen geplantem und ungeplantem Kaiserschnitt vor, sodass zu dieser Ausprägung des Endpunkts Geburtsmodus keine Informationen extrahiert wurden. In 1 weiteren Studie werden die ungeplanten Kaiserschnitte während der Geburt berichtet, aus deren Anzahl die Berechnung der Anzahl der vaginalen Geburten ohne Unterscheidung zwischen Spontangeburt und vaginal-operativer Geburt möglich ist.

Für die klinisch wichtigen, aber für die Empfehlung nicht entscheidenden Endpunkte liegen Informationen zu 2 Endpunkten vor: Episiotomie und Wehenmittel. Für die Endpunkte Amniotomie, Kristellerhilfe, Infektionen, leichte, mittelgradige oder fortgeschrittene Acidose, Ikterus und Anämie werden keine Informationen berichtet.

In keiner der Studien wurde eine Unterscheidung zwischen primären und sekundären Endpunkten vorgenommen.

Tabelle 13: Matrix der kritischen und klinisch wichtigen Endpunkte zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Kritische Endpunkte													Klinisch wichtige Endpunkte									
	Maternal									Neonatal				Maternal				Neonatal					
	Mortalität	Geburtsmodus			Höhergradige Dammmrisse (III / IV°)	Hoher mütterlicher Blutverlust ≥ 1000 ml	Stillbeziehung bei initialem Stillwunsch	Teilweise / vollständige Plazentaretention	Geburtserleben	Subjektive Gesundheitseinschätzung	Mortalität	Apgar-Score nach 5 Minuten < 6	Schwere metabolische Acidose	Basendefizit ≥ -16 mmol/l	Verlegung auf die NICU	Episiotomie	Wehenmittel	Amniotomie	Kristellerhilfe	Infektionen	Leichte, mittelgradige, fortgeschrittene Acidose	Ikterus	Anämie
Spontangeburt		Vaginal-operative Geburt	Kaiserschnitt																				
Belghiti 2013	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	■ ^a	-	-	-	-	-	-	-
Chuilon 2016	-	-	-	-	-	-	-	-	-	-	-	-	-	-	■	-	-	-	-	-	-	-	-
Coulm 2012	-	○	■	- ^b	-	-	-	-	-	-	-	-	-	-	■ ^c	-	-	-	-	-	-	-	-
Finnström 2006	-	-	-	-	-	-	-	-	-	■ ^d	-	-	-	-	-	-	-	-	-	-	-	-	-
Karalis 2016	-	-	-	-	-	-	-	-	-	■ ^e	-	■	-	-	-	-	-	-	-	-	-	-	-
Wehberg 2018	-	○ ^f	■ ^f	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

● primärer Endpunkt; ■ weitere / sekundäre Endpunkte; ○ aus berichteten Daten ableitbar; – der Endpunkt wurde nicht erhoben
 a: Berichtet wird die Gabe von Oxytocin während der Geburt bei Gebärenden mit französischer Standardbetreuung.
 b: Berichtet wird die Gesamtanzahl der Kaiserschnitte ohne Differenzierungsmöglichkeit zwischen geplantem und ungeplantem Kaiserschnitt, sodass keine Informationen extrahiert wurden.
 c: Berichtet werden Ereignisse für Frauen mit einer Spontangeburt (ohne vaginal-operative Geburt).
 d: Berichtet werden neonatale Todesfälle (1.–27. Lebenstag).
 e: Berichtet werden intrapartale sowie frühe neonatale Todesfälle (1.–7. Lebenstag).
 f: Berichtet werden die ungeplanten Kaiserschnitte während der Geburt. Daraus ist die Berechnung der Anzahl der vaginalen Geburten ohne Unterscheidung zwischen Spontangeburt und vaginal-operativer Geburt möglich.
 NICU: Neugeborenen-Intensivstation

7.2 Einschätzung des Verzerrungspotenzials

7.2.1 Einschätzung des Verzerrungspotenzials auf Studienebene

Für die Bewertung lagen 6 Beobachtungsstudien vor. Die Einschätzung des Verzerrungspotenzials auf Studienebene ist in Tabelle 14 dargestellt.

Bei allen 6 Studien handelte es sich um nicht randomisierte Kohortenstudien mit einer zeitlich parallelen Kontrollgruppe und adäquater Confounderkontrolle auf der Basis einer prospektiv geplanten Datenerfassung in Geburtsregistern. Die Vergleichbarkeit der Gruppen beziehungsweise die adäquate Berücksichtigung der prognostisch relevanten Faktoren wurde für 4 Studien (Belghiti 2013, Chuilon 2016, Coulm 2012, Wehberg 2018) mit ja und für 2 Studien (Finnström 2006, Karalis 2016) mit unklar bewertet.

In Finnström 2006 liegen Ergebnisse nach einer Confounderadjustierung für das maternale Alter, Parität, Rauchen, BMI und SSW vor. Da die Daten aller lebenden Einlingsgeburten ausgewertet wurden, wurden für die Auswertung des Endpunkts neonatale Mortalität Entwicklungsstörungen der Feten und maternale Grund- oder schwangerschaftsbedingte Erkrankungen (z. B. Plazentastörungen) nicht ausreichend berücksichtigt. In Karalis 2016 wurde für das maternale Alter und Parität adjustiert, und Neugeborene mit schweren angeborenen Fehlbildungen und einem Geburtsgewicht unter 2500 g wurden ausgeschlossen. Für die Auswertung der neonatalen Endpunkte neonatale Mortalität und schwere metabolische Acidose wurden maternale Erkrankungen und Entwicklungsstörungen der Feten nicht ausreichend berücksichtigt. Frauen haben in der Regel eine Wahlmöglichkeit, welches Krankenhaus beziehungsweise welche Geburtsklinik sie zur Geburt aufsuchen, und können dies anhand verschiedener Merkmale, z. B. Größe, Ausstattung oder Lage, selbst bestimmen. Frauen mit Risiken oder speziellem Behandlungsbedarf werden in der Regel aufgrund ihrer Charakteristika in eine bestimmte Geburtsklinik überwiesen. Eine vollständige Adjustierung für alle potenziellen maternalen und fetalen Risikofaktoren war, auch aufgrund der teilweise fehlenden Erfassung der Informationen im Geburtsregister, nicht möglich.

In 2 Studien (Belghiti 2013, Coulm 2012) auf der Basis der französischen Prinataldatenerhebung werden Daten für Erst- und Mehrgebärende mit einem geringen Risiko und einer termingerechten Geburt eines lebenden Einlings in Schädellage berichtet. Ausgeschlossen wurden alle Frauen mit schweren maternalen Erkrankungen (Bluthochdruck, Diabetes Typ 1, chronischem Organversagen, Thrombophilie, Lupus erythematodes, Antiphospholipid-Syndrom, Epilepsie, chronischen Infektionen mit Hepatitis B / C / HIV), einer ungünstigen geburtshilflichen Anamnese (vorheriger Kaiserschnitt, Schwangerschaftscholestase, Präeklampsie oder Totgeburt), Erkrankungen in der aktuellen Schwangerschaft (schwere Präeklampsie, insulinpflichtiger Gestationsdiabetes, Rhesus-Inkompatibilität) oder intrauteriner Wachstumsretardierung < 5. Perzentil. Für alle untersuchten Endpunkte erfolgte eine Adjustierung sowohl für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad) als auch neonatale Charakteristika (SSW, Geburtsgewicht).

Die 3. Auswertung der französischen Perinatalerhebung (Chuilon 2016) wertet für alle Erst- und Mehrgebärende ab der 22. SSW oder mit einem Neugeborenen mit einem Geburtsgewicht > 500 g mit einer vaginalen Geburt die Ergebnisse zur Episiotomie aus. Ausgeschlossen wurden Frauen mit einer Mehrlingsschwangerschaft, einem intrauterin verstorbenen Kind, einer Abruption oder bei fehlenden Informationen zur Episiotomie. Es erfolgte eine Adjustierung sowohl für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad, vorheriger Kaiserschnitt) und neonatale Charakteristika (Lage, SSW, Geburtsgewicht) als auch für Charakteristika der Geburt (PDA, Wehenmittelgabe, betreuendes Fachpersonal, Geburtsmodus) und der Geburtsklinik (Größe, Versorgungsstufe, Ausstattung).

In Wehberg 2018 werden Ergebnisse für Erstgebärende ab der vollendeten 37. SSW mit einem geringen Risiko, die ein Kind in Schädellage mit einer spontanen Wehentätigkeit gebären, berichtet. Es erfolgte eine Adjustierung sowohl für maternale Charakteristika (Alter, BMI, Rauchen, Größe, Diabetes, Bluthochdruck, Eklampsie, Kopf-Becken-Missverhältnis, Plazenta Praevia, vorzeitige Plazentalösung und vorzeitiger Blasensprung), neonatale Charakteristika (SSW, Geschlecht, Fetal Distress, Makrosomie), geburtsbezogene Charakteristika (Geburtseinleitung) als auch systembezogene Charakteristika (Geburt am Wochenende oder während der Nacht).

Die Gebärenden und das behandelnde Personal waren nicht verblindet, da Gebärende und behandelndes Personal in der Regel über Größe und Versorgungsstufe der Geburtsklinik informiert sind. Da hier die fehlende Verblindung Teil der Exposition war, resultiert daraus keine Verzerrung.

Die ergebnisunabhängige Berichterstattung wurde für 5 Studien als unklar bewertet, da keine Information zu den geplanten Endpunkten vorliegt, da weder ein Studienprotokoll noch ein Studienregistereintrag verfügbar waren. In 1 Studie lag keine ergebnisunabhängige Berichterstattung vor, da ausschließlich signifikante Ergebnisse berichtet werden.

In allen Studien lagen weitere Aspekte vor, die das Verzerrungspotenzial beeinflussen können. Diese beinhalten einen strukturellen Unterschied in der materiellen und personellen Ausstattung der Geburtskliniken, die nicht die Erfahrung/Expertise der Geburtshelfer widerspiegeln, und die schlechtere Erreichbarkeit von Kliniken in peripher gelegenen Gebieten.

In der Folge wurde das Verzerrungspotenzial auf Studienebene für alle Studien als hoch bewertet.

Tabelle 14: Verzerrungspotenzial auf Studienebene zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung prognostisch relevanter Faktoren	Verblindung				Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Studienebene
			Gebärende	Behandelnde Personen	Ergebnisunabhängige Berichterstattung			
Belghiti 2013	ja	ja	nein	nein	unklar	nein	hoch ^{a, b}	
Chuilon 2016	ja	ja	nein	nein	unklar	nein	hoch ^{a, b}	
Coulm 2012	ja	ja	nein	nein	unklar	nein	hoch ^{a, b}	
Finnström 2006	ja	unklar	nein	nein	nein	nein	hoch ^{a, c, d}	
Karalis 2016	ja	unklar	nein	nein	unklar	nein	hoch ^{a, b, c}	
Wehberg 2018	ja	ja	nein	nein	unklar	nein	hoch ^{a, b, e}	
a: fehlende Verblindung von Gebärenden und geburtshilflichem Fachpersonal (keine Abwertung da Teil der Exposition)								
b: unklare ergebnisunabhängige Berichterstattung aufgrund eines fehlenden Studienprotokolls								
c: eingeschränkt adäquate Confounderadjustierung und fehlende Informationen zu Baseline-Charakteristika in Expositions- und Referenzgruppen								
d: fehlende ergebnisunabhängige Berichterstattung aufgrund eines fehlenden Studienprotokolls und selektiver Berichterstattung: Vergleich von Kliniken mit und ohne Kinderklinik wird für die Vergleiche zur Größe des Einzugsgebietes, nicht aber für die Größe der Kliniken berichtet								
e: Unterschiedliche Erfassung bestimmter Risikofaktoren ohne prospektiv festgelegte präzise Definition kann nicht ausgeschlossen werden, da diese erst nach der Geburt erfasst wurden.								

7.2.2 Einschätzung des Verzerrungspotenzials auf Endpunktebene

Geburtsmodus

Das Verzerrungspotenzial für den Geburtsmodus basiert auf dem Verzerrungspotenzial auf Studienebene und wurde in den 2 Studien mit hoch bewertet (siehe Tabelle 31 im Anhang).

Neonatale Mortalität

Das Verzerrungspotenzial für die neonatale Mortalität basiert auf dem Verzerrungspotenzial auf Studienebene und wurde in den 2 Studien mit hoch bewertet (siehe Tabelle 32 im Anhang). Das Verzerrungspotenzial wird zusätzlich durch den Ausschluss von 324 (0,12 %) der 267 390 Gebärenden beziehungsweise Neugeborenen aus der Analyse aufgrund fehlender Werte beeinflusst. Ein Einfluss auf das Ergebnis für die seltenen Ereignisse der intrapartalen und frühen neonatalen Todesfälle, die bei insgesamt 709 (0,26 %) und 54 (0,02 %) aller Neugeborenen auftraten, kann nicht ausgeschlossen werden.

Schwere metabolische Acidose (pH < 7,0)

Das Verzerrungspotenzial für das Auftreten einer schweren metabolischen Acidose basiert auf dem Verzerrungspotenzial auf Studienebene und wurde in der 1 Studie mit hoch bewertet (siehe Tabelle 33 im Anhang). Das Verzerrungspotenzial wird zusätzlich durch den Ausschluss von 324 (0,12 %) der 267 390 Gebärenden beziehungsweise Neugeborenen aus der Analyse aufgrund fehlender Werte beeinflusst. Ein Einfluss auf das Ergebnis für dieses seltene Ereignis, das bei insgesamt 874 (0,37 %) aller Neugeborenen auftrat, kann nicht ausgeschlossen werden.

Episiotomie

Das Verzerrungspotenzial für das Auftreten von Episiotomien basiert auf dem Verzerrungspotenzial auf Studienebene und wurde in den 2 Studien mit hoch bewertet (siehe Tabelle 34 im Anhang). Informationen zur Geburt basieren auf den medizinischen Unterlagen, eine Beeinflussung dieses subjektiv beeinflussbaren Endpunktes kann nicht sicher ausgeschlossen werden.

Wehenmittel

Das Verzerrungspotenzial für die Gabe von Wehenmitteln basiert auf dem Verzerrungspotenzial auf Studienebene und wurde in den 2 Studien mit hoch bewertet (siehe Tabelle 35 im Anhang). Informationen zur Geburt basieren auf den medizinischen Unterlagen, eine Beeinflussung dieses subjektiv beeinflussbaren Endpunktes kann nicht sicher ausgeschlossen werden.

7.3 Ergebnisse zu kritischen Endpunkten

Die nachfolgende Tabelle 15 gibt einen Überblick über die Ergebnisse der Evidenz zu den kritischen Endpunkten für die eingeschlossenen Studien. Insgesamt wurden aufgrund der hohen Fallzahlen im Vergleich zu den vorherigen Rapid Reports verhältnismäßig viele statistisch signifikante Zusammenhänge nachgewiesen, welche aufgrund des hohen Verzerrungspotenzials der eingeschlossenen Studien kritisch zu interpretieren sind.

Fragestellung dieses Rapid Reports ist der Vergleich größerer und kleinerer Erfahrung / Expertise. Deshalb werden die Ergebnisse aus den primär interessierenden Vergleichen zwischen den größten und kleinsten Ausprägungen von Erfahrung / Expertise interpretiert, soweit diese auf Deutschland übertragbar sind. Erfahrung / Expertise wird anhand der Geburtenzahl pro Jahr oder der Versorgungsstufe der Geburtsklinik beurteilt. Da in Deutschland 2017 nur sehr wenige Krankenhäuser weniger als 100 Geburten oder mehr als 4000 Geburten pro Jahr aufwiesen [61], wurden Interventionsgruppen mit > 4000 Geburten pro Jahr und Referenzgruppen mit < 100 Geburten pro Jahr nicht als primär interessierende Analyse herangezogen.

Tabelle 15: Zusammenfassung der Ergebnisse zu den kritischen Endpunkten zur Erfahrung / Expertise durch systemimmanente Faktoren

Endpunkt	Ergebnisse für die primär interessierenden Vergleiche der größten und kleinsten Ausprägung von Erfahrung / Expertise (Studie)
Geburtsmodus	<ul style="list-style-type: none"> ▪ Geburtenzahl pro Jahr <ul style="list-style-type: none"> ▫ Spontangeburt: keine Effektschätzer berichtet ▫ Vaginal-operative Geburt: <ul style="list-style-type: none"> - ≥ 3000 vs. < 1000: aOR: 1,47; 98,3 %-KI: [1,03; 2,09] (Coulm 2012) ▫ Vaginale Geburt: <ul style="list-style-type: none"> - > 3000 vs. 1100–3000: aOR: 1,22; 97,5 %-KI: [1,14; 1,30] (Wehberg 2018) - 1100–3000 vs. < 1100: aOR: 1,02; 97,5 %-KI: [0,91; 1,15] (Wehberg 2018) ▫ Kaiserschnitt: <ul style="list-style-type: none"> - > 3000 vs. 1100–3000: aOR: 0,82; 98,3 %-KI: [0,77; 0,88] (Wehberg 2018) - 1100–3000 vs. < 1100: aOR: 0,98; 98,3 %-KI: [0,86; 1,12] (Wehberg 2018) ▪ Versorgungstufe der Geburtsklinik <ul style="list-style-type: none"> ▫ vaginal-operative Geburt: <ul style="list-style-type: none"> - Level 3 vs. Level 1: aOR: 1,30; 98,3 %-KI: [0,95; 1,77] (Coulm 2012)
neonatale Mortalität	<ul style="list-style-type: none"> ▪ Geburtenzahl pro Jahr: <ul style="list-style-type: none"> ▫ neonatale Todesfälle (Tag 1–27): <ul style="list-style-type: none"> - ≥ 1000–2499 vs. < 500: aOR: 1,19; 95 %-KI: [0,90; 1,59] (Finnström 2006) ▫ frühe neonatale Todesfälle (Tag 1–7): <ul style="list-style-type: none"> - Universitätskrankenhaus vs. ≤ 999: aOR: 0,47; 95 %-KI: [0,22; 1,03] (Karalis 2016) ▫ intrapartale Todesfälle: <ul style="list-style-type: none"> - Universitätskrankenhaus vs. ≤ 999: aOR: 0,84; 95 %-KI: [0,66; 1,05] (Karalis 2016)
Schwere metabolische Acidose (pH $< 7,0$)	<ul style="list-style-type: none"> ▪ Universitätskrankenhaus vs. ≤ 999: aOR: 1,43; 95 %-KI: [1,14; 1,79] (Karalis 2016)
aOR: adjustiertes Odds Ratio; KI: Konfidenzintervall; vs.: versus	

7.3.1 Geburtsmodus

Insgesamt lagen zum Endpunkt Geburtsmodus confounderadjustierte Ergebnisse aus 2 Studien vor, wobei in 2 Studien der Einfluss der Anzahl der Geburten der Geburtsklinik pro Jahr und in 1 Studie zusätzlich der Einfluss der Versorgungsstufe der Geburtsklinik auf den Geburtsmodus berichtet wurde (siehe Tabelle 51 und Tabelle 52 im Anhang).

In Coulm 2012 wurden Ergebnisse für das Ereignis vaginal-operative Geburt und in Wehberg 2018 für die 2 Ereignisse vaginale Geburt und ungeplanter Kaiserschnitt berichtet. Die Angabe der vaginalen Geburten beinhaltet keine Differenzierung zwischen Spontangeburt und vaginal-operativer Geburt, sondern fasst beide Ereignisse zusammen. Eine Nachberechnung der Ergebnisse auf Grundlage der Ereigniszahlen erfolgte nicht, da eine Confounderadjustierung ohne Vorliegen der Individualdaten nicht möglich ist.

Die 3 Ereignisse Spontangeburt, vaginal-operative Geburt und Kaiserschnitt sind 3 mögliche Ausprägungen des nominalen Endpunkts Geburtsmodus. Diese Ausprägungen wurden wie

3 separate binäre Endpunkte ausgewertet und ihre gegenseitige Abhängigkeit wurde durch die Adjustierung des Signifikanzniveaus der Analysen für multiples Testen mit der Bonferroni-Methode und der Berechnung der 98,3 %-KI berücksichtigt. Die Ereignisse vaginale Geburt und Kaiserschnitt beschreiben 2 mögliche Ausprägungen des nominalen Endpunkts Geburtsmodus, sodass die Anwendung der Bonferroni-Methode für multiples Testen in der Verwendung von 97,5 %-KI für das Ereignis vaginale Geburt resultierte. Da eine endpunktspezifische Bewertung der Qualität der Evidenz geplant ist, werden Ergebnisse zu den 3 Endpunkten Spontangeburt, vaginal-operative Geburt und Kaiserschnitt konsistent mit ihren 98,3 %-KI und Ergebnisse zum Endpunkt vaginale Geburt konsistent mit ihren 97,5 %-KI berichtet.

Spontangeburt

Informationen zum Einfluss der Anzahl der Geburten der Geburtsklinik pro Jahr auf die Anzahl von Spontangeburt werden in 1 Studie (Coulm 2012) berichtet (siehe Tabelle 51 und Tabelle 52 im Anhang).

Für Coulm 2012 konnte die absolute Anzahl von Spontangeburt aus den Angaben zur Episiotomie bei Spontangeburt berechnet werden, aufgrund der fehlenden exakten Angaben zur Anzahl der ausgewerteten Gebärenden in den Vergleichsgruppen bleiben die relativen Häufigkeiten unklar. Es wurden keine confounderadjustierten Ergebnisse für diese Vergleiche berichtet, sodass eine Zusammenfassung des Einflusses systemimmanenter Faktoren auf den Endpunkt Spontangeburt nicht möglich ist.

Vaginal-operative Geburt

Informationen zum Einfluss der Anzahl der Geburten der Geburtskliniken pro Jahr und der Versorgungsstufe auf die Anzahl von vaginal-operativen Geburten werden in 1 Studie (Coulm 2012) berichtet (siehe Tabelle 51 und Tabelle 52 im Anhang).

In Coulm 2012 wiesen insgesamt 16,5 % der Erst- und Mehrgebärenden in der Gruppe der großen Geburtskliniken mit ≥ 3000 Geburten pro Jahr eine vaginal-operative Geburt verglichen mit 16,4 %, 14,8 % und 12,7 % der Gebärenden in den Gruppen der kleineren Geburtskliniken mit 2000–2999, 1000–1999 und < 1000 Geburten pro Jahr auf. Aufgrund der fehlenden Bezugsgröße ist eine Berechnung der absoluten Häufigkeiten nicht möglich. Die Effektschätzungen steigen mit wachsendem Abstand zwischen den Expositionen, da mit abnehmender Größe der Geburtskliniken die Häufigkeit einer vaginal-operativen Geburt konsistent abnimmt. In spezialisierten Tertiärkrankenhäusern mit einer Neugeborenen-Intensivstation (Versorgungsstufe Level 3) wiesen 15,0 % der Gebärenden eine vaginal-operative Geburt auf, verglichen mit 13,0 %, 16,1 % und 15,9 % der Gebärenden in den Vergleichsgruppen in Geburtskliniken mit einer geringeren Versorgungsstufe (Level 2b, 2a und 1).

In Coulm 2012 konnte der primär interessierende Vergleich zwischen Geburtskliniken mit einer größeren Geburtenzahl (≥ 3000 Geburten pro Jahr) und einer geringeren Geburtenzahl (< 1000 Geburten pro Jahr) eine statistisch signifikant höhere Anzahl vaginal-operativer

Geburten in Geburtskliniken mit höheren Geburtenzahlen nachweisen (OR: 1,47; 98,3 %-KI: [1,03; 2,09]). Da Gebärende mit einem erhöhten Risiko ausgeschlossen wurden und eine adäquate Confounderadjustierung erfolgte, ist davon auszugehen, dass ein möglicher Selektionsbias angemessen in der Auswertung berücksichtigt wurde und die Ergebnisse interpretierbar sind. Aufgrund der fehlenden Informationen zu den weiteren Ausprägungen des Endpunkts Spontangeburt und Kaiserschnitt kann keine Aussage darüber getroffen werden, auf welche dieser Ausprägungen das abnehmende Risiko einer vaginal-operativen Geburt Einfluss nehmen würde. Der primär interessierende Vergleich zwischen den Versorgungsstufen Level 3 und Level 1 konnte keinen statistisch signifikanten Unterschied in Abhängigkeit von der Versorgungsstufe nachweisen (OR: 1,30; 98,3 %-KI: [0,95; 1,77]).

Vaginale Geburt

Informationen zum Einfluss der Anzahl der Geburten der Geburtskliniken pro Jahr auf die Anzahl von vaginalen Geburten werden in 1 Studie (Wehberg 2018) berichtet (siehe Tabelle 51 im Anhang), in welcher die Anzahl vaginaler Geburten und Kaiserschnitte bei Erstgebärenden berichtet wird.

Insgesamt 65 304 (88,2 %) der eingeschlossenen 74 002 Erstgebärenden entbanden vaginal: 88,6 % der Gebärenden in der Gruppe der großen Geburtskliniken mit > 3000 Geburten pro Jahr verglichen mit 87,8 % und 87,5 % der Gebärenden in den mittelgroßen und kleinen Geburtskliniken mit 1100–3000 und < 1100 Geburten pro Jahr.

Es werden Vergleiche zwischen großen Geburtskliniken mit > 3000 Geburten pro Jahr und mittelgroßen Geburtskliniken mit 1100–3000 Geburten sowie zwischen mittelgroßen Geburtskliniken und kleinen Geburtskliniken mit < 1100 Geburten pro Jahr berichtet. Da Gebärende mit einem erhöhten Risiko ausgeschlossen wurden und eine adäquate Confounderadjustierung erfolgte, ist davon auszugehen, dass ein möglicher Selektionsbias angemessen in der Auswertung berücksichtigt wurde und die Ergebnisse interpretierbar sind. Es konnte ein statistisch signifikanter Unterschied mit einer höheren Anzahl vaginaler Geburten bei Erstgebärenden in der Gruppe der großen Geburtskliniken mit > 3000 Geburten pro Jahr im Vergleich zu mittelgroßen Geburtskliniken mit 1100–3000 Geburten pro Jahr (OR: 1,22; 97,5 %-KI: [1,14; 1,30]) nachgewiesen werden. Zwischen den mittelgroßen und kleinen Geburtskliniken mit < 1100 Geburten pro Jahr konnte kein statistisch signifikanter Unterschied nachgewiesen werden (OR: 1,02; 97,5 %-KI: [0,91; 1,15]).

Kaiserschnitte

Informationen zum Einfluss der Anzahl der Geburten pro Jahr auf die Anzahl von Kaiserschnitten bei Erstgebärenden werden in 1 Studie (Wehberg 2018) berichtet (siehe Tabelle 51 im Anhang).

Insgesamt 8698 (11,8 %) der eingeschlossenen 74 002 Erstgebärenden wiesen einen ungeplanten Kaiserschnitt auf: 11,4 % der Gebärenden in der Gruppe der großen Geburtskliniken mit > 3000 Geburten pro Jahr verglichen mit 12,2 % und 12,5 % der

Gebärenden in den Gruppen der mittelgroßen und kleineren Geburtskliniken mit 1100–3000 und < 1100 Geburten pro Jahr.

Es werden Vergleiche zwischen großen Geburtskliniken mit > 3000 Geburten pro Jahr und mittelgroßen Geburtskliniken mit 1100–3000 Geburten sowie zwischen mittelgroßen Geburtskliniken und kleinen Geburtskliniken mit < 1100 Geburten pro Jahr berichtet. Da Gebärende mit einem erhöhten Risiko ausgeschlossen wurden und eine adäquate Confounderadjustierung erfolgte, ist davon auszugehen, dass ein möglicher Selektionsbias angemessen in der Auswertung berücksichtigt wurde und die Ergebnisse interpretierbar sind. Es konnte für die ausgewerteten Erstgebärenden ein statistisch signifikanter Unterschied mit einer geringeren Anzahl von ungeplanten Kaiserschnitten in großen Geburtskliniken mit > 3000 Geburten pro Jahr im Vergleich zu mittelgroßen Geburtskliniken mit 1100 bis 3000 Geburten pro Jahr (OR: 0,82; 98,3 %-KI: [0,77; 0,88]) nachgewiesen werden. Zwischen mittleren und kleinen Geburtskliniken mit < 1100 Geburten pro Jahr konnte kein statistisch signifikanter Unterschied nachgewiesen werden (OR: 0,98; 98,3 %-KI: [0,86; 1,12]).

7.3.2 Neonatale Mortalität

Informationen zum Einfluss der Anzahl der Geburten der Geburtskliniken pro Jahr auf die Anzahl neonataler Todesfälle werden in 2 Studien (Finnström 2006, Karalis 2016) berichtet (siehe Tabelle 53 im Anhang), wobei 1 dieser Studien (Finnström 2006) alle lebend geborenen Einlinge ohne Beschränkung des Gestationsalters in die Auswertung einbezog. Während bei Finnström 2006 alle neonatalen Todesfälle innerhalb von 27 Tagen nach der Geburt in die Auswertung eingehen, beschränkt sich Karalis 2016 auf die Auswertung der intrapartalen und der frühen neonatalen Todesfälle innerhalb von 7 Tagen nach der Geburt.

Insgesamt sind die Mortalitätsraten für Neugeborene von Frauen mit einem geringen Risiko für Komplikationen sehr niedrig und liegen in Karalis 2016 bei 0,02 % (2 von 10 000 Neugeborenen). In Finnström 2006 werden keine absoluten Angaben zur Anzahl der neonatalen Mortalität berichtet, es liegen ausschließlich Angaben zu Effektschätzern vor.

In Karalis 2016 verstarben insgesamt 54 (0,02 %) der eingeschlossenen 267 390 Neugeborenen innerhalb von 7 Tagen nach der Geburt: 0,02 % der im Universitätskrankenhaus geborenen Neugeborenen verglichen mit 0,02 %, 0,03 % und 0,03 % der in außeruniversitären Geburtskliniken geborenen Neugeborenen mit ≥ 2000 , 1000–1999 und ≤ 999 Geburten pro Jahr. Zudem lagen Informationen zu den intrapartalen Todesfällen vor. Es verstarben insgesamt 709 (0,27 %) der eingeschlossenen 267 390 Neugeborenen während der Geburt: 0,24 % der Neugeborenen im Universitätskrankenhaus verglichen mit 0,27 %, 0,30 % und 0,27 % der in außeruniversitären Geburtskliniken geborenen Neugeborenen mit ≥ 2000 , 1000–1999 und ≤ 999 Geburten pro Jahr. Die Effektschätzungen nehmen mit wachsendem Abstand zwischen Expositionen zu, da mit abnehmender Größe der Geburtskliniken die neonatale Mortalität zunimmt, ohne dass ein signifikanter Unterschied nachgewiesen werden konnte.

In Finnström 2006 konnte für die primär interessierende Analyse zwischen Geburtskliniken mit einer größeren Geburtenzahl (≥ 1000 –2499 Geburten pro Jahr) und einer geringeren Geburtenzahl (< 500 Geburten pro Jahr) kein statistisch signifikanter Zusammenhang zwischen Geburtenzahl und der neonatalen Mortalität nachgewiesen werden (OR: 1,19; 95 %-KI: [0,90; 1,59]). Die Autorinnen und Autoren berichten, dass aus den Einzugsbereichen der kleinsten Geburtskliniken mit < 500 Geburten im Jahr 20 % der Gebärenden in größere spezialisierte Geburtskliniken überwiesen wurden, und verglichen die Mortalität von Neugeborenen aus den Einzugsgebieten größerer und kleinerer Geburtskliniken. Es wurde keine höhere Mortalität für Neugeborene von Gebärenden im Einzugsgebiet kleinerer Geburtskliniken nachgewiesen.

In Karalis 2016 konnte für keinen der Vergleiche zwischen universitären und nicht universitären Krankenhäusern ein statistisch signifikanter Unterschied hinsichtlich der neonatalen Mortalität nachgewiesen werden. So zeigte auch der primäre Vergleich zwischen Universitätskrankenhäusern und nicht universitären Krankenhäusern mit ≤ 999 Geburten pro Jahr weder für die frühe neonatale Mortalität (OR: 0,47; 95 %-KI: [0,22; 1,03]) noch für das Auftreten intrapartaler Todesfälle (OR: 0,84; 95 %-KI: [0,66; 1,05]) einen statistisch signifikanten Unterschied.

7.3.3 Apgar-Score nach 5 Minuten < 6

Für den Apgar-Score nach 5 Minuten von < 6 liegen ausschließlich ergänzende Informationen in 2 Studien vor (siehe Tabelle 54 im Anhang).

7.3.4 Schwere metabolische Acidose (pH $< 7,0$)

Informationen zum Einfluss der Anzahl der Geburten der Krankenhäuser pro Jahr auf die Anzahl von Neugeborenen mit einer schweren metabolischen Acidose (pH $< 7,0$) werden in 1 Studie (Karalis 2016) berichtet (siehe Tabelle 55 im Anhang).

Insgesamt 874 (0,3 %) der eingeschlossenen 267 390 Neugeborenen litten unter einer schweren metabolischen Acidose: 0,39 % der im Universitätskrankenhaus geborenen Neugeborenen verglichen mit 0,27 %, 0,35 % und 0,25 % der in außeruniversitären Krankenhäusern mit ≥ 2000 , 1000–1999 und ≤ 999 Geburten pro Jahr geborenen Neugeborenen.

Für den primär interessierenden Vergleich konnte ein statistisch signifikanter Unterschied mit einer höheren Anzahl Neugeborener mit einer schweren metabolischen Acidose in der Gruppe der Universitätskrankenhäuser im Vergleich zu kleineren außeruniversitären Geburtskliniken mit ≤ 999 Geburten im Jahr (OR: 1,43; 95 %-KI: [1,14; 1,79]) nachgewiesen werden.

7.3.5 Verlegung auf die Neugeborenen-Intensivstation

Für die Verlegung auf die Neugeborenen-Intensivstation liegen ausschließlich ergänzende Informationen in 1 Studie vor (siehe Tabelle 56 im Anhang).

7.4 Ergebnisse zu wichtigen Endpunkten

Die nachfolgende Tabelle 16 gibt einen Überblick über die Ergebnisse der Evidenz zu den wichtigen Endpunkten für die eingeschlossenen Studien.

Tabelle 16: Zusammenfassung der Ergebnisse zu den wichtigen Endpunkten zur Erfahrung / Expertise durch systemimmanente Faktoren

Endpunkt	Ergebnisse für die primär interessierenden Vergleiche der größten und kleinsten Ausprägung von Erfahrung / Expertise (Studie, zusätzliche Information)
Episiotomie	<ul style="list-style-type: none"> ▪ Geburtenzahl pro Jahr <ul style="list-style-type: none"> ▫ ≥ 3000 vs. < 1000: aOR: 0,80; 95 %-KI: [0,57; 1,12] (Coulm 2012, Frauen mit Spontangeburt) ▪ Versorgungsstufe der Geburtsklinik <ul style="list-style-type: none"> ▫ Level 3 vs. Level 1: aOR: 0,96; 95 %-KI: [0,70; 1,31] (Coulm 2012, Frauen mit Spontangeburt) ▫ Level 3 vs. Level 1: aRR: 1,11; 95 %-KI: [0,91; 1,36] (Chuilon 2016, Frauen mit vaginaler Geburt)
Wehenmittel	<ul style="list-style-type: none"> ▪ Geburtenzahl pro Jahr <ul style="list-style-type: none"> ▫ Wehenmittelgabe während der Geburt: <ul style="list-style-type: none"> - ≥ 4000 vs. < 1000: aOR: 0,67; 95 %-KI: [0,56; 0,83] (Belghiti 2013) ▪ Versorgungsstufe der Geburtsklinik <ul style="list-style-type: none"> ▫ Wehenmittelgabe während der Geburt: <ul style="list-style-type: none"> - Level 3 vs. Level 1: aOR: 0,63; 95 %-KI: [0,51; 0,77] (Belghiti 2013, Frauen mit einer spontanen Wehentätigkeit und PDA)
aOR: adjustiertes Odds Ratio; aRR: adjustiertes relatives Risiko; KI: Konfidenzintervall; PDA: Periduralanästhesie; vs.: versus	

7.4.1 Episiotomie

Informationen zum Einfluss der Anzahl der Geburten pro Jahr oder der Versorgungsstufe der Geburtskliniken auf die Anzahl von Episiotomien werden in jeweils 2 Studien (Coulm 2012, Chuilon 2016) berichtet (siehe Tabelle 57 und Tabelle 58 im Anhang).

In Coulm 2012 wurde bei 1395 (19,6 %) der 7122 Frauen mit einer Spontangeburt eine Episiotomie durchgeführt: 17,0 % in der Gruppe der großen Geburtskliniken mit ≥ 3000 Geburten pro Jahr verglichen mit 19,2 %, 21,0 % und 19,9 % in kleineren Geburtskliniken mit 2000–2999, 1000–1999 und < 1000 Geburten pro Jahr. Aufgrund der aus den relativen Häufigkeiten berechneten absoluten Anzahlen wurde in der Auswertung der Versorgungsstufe eine abweichende Anzahl von 1396 Gebärenden berechnet, bei denen eine Episiotomie durchgeführt wurde. Von diesen 1396 (19,6 %) der 7122 Frauen mit einer Spontangeburt wurden 16,6 % der Episiotomien in der Gruppe der Geburtskliniken der höchsten Versorgungsstufe durchgeführt, verglichen mit 20,6 %, 21,4 % und 19,5 % in Geburtskliniken mit einer niedrigeren Versorgungsstufe (Level 2b, 2a und 1).

Chuilon 2016 greift auf dieselbe Datenbasis wie Coulm 2012 (französische Perinataldatenerhebung 2010) zurück und berichtet die Häufigkeit von Episiotomien bei Erst-

und Mehrgebärenden mit einer vaginalen Geburt ohne Beschränkung auf Frauen mit einem geringen Risiko für Komplikationen. Alle Vergleiche beziehen sich auf Geburtskliniken mit ≥ 4000 Geburten pro Jahr. Da in Deutschland 2017 nur sehr wenige Krankenhäuser mehr als 4000 Geburten pro Jahr aufwiesen [61] und damit die Übertragbarkeit der berichteten Ergebnisse eingeschränkt ist, werden die Ergebnisse aus Chuilon 2016 hier ergänzend berichtet. Bei insgesamt 3020 (26,8 %) der eingeschlossenen Frauen wurde eine Episiotomie durchgeführt: bei 2096 (44,7 %) der Erstgebärenden und 924 (14,3 %) der Mehrgebärenden. In der Gruppe der 4689 ausgewerteten Erstgebärenden wurde bei 41,6 % der Gebärenden in der Gruppe der sehr großen Geburtskliniken mit ≥ 4000 Geburten pro Jahr eine Episiotomie durchgeführt verglichen mit 43,1 %, 44,9 %, 44,6 % und 46,7 % in kleineren Geburtskliniken mit 3000–3999, 2000–2999, 1000–1999 und < 1000 Geburten pro Jahr. In der Gruppe der 6453 ausgewerteten Mehrgebärenden wurde bei 12,4 % der Gebärenden in der Gruppe der sehr großen Geburtskliniken mit ≥ 4000 Geburten pro Jahr eine Episiotomie durchgeführt verglichen mit 12,3 %, 14,7 %, 15,1 % und 14,2 % in kleineren Geburtskliniken mit 3000–3999, 2000–2999, 1000–1999 und < 1000 Geburten pro Jahr.

Zudem liegen aus Chuilon 2016 Informationen zur Versorgungsstufe der Geburtskliniken vor. Von den 4689 ausgewerteten Erstgebärenden mit einer vaginalen Geburt wurde bei 2096 (19,6 %) eine Episiotomie durchgeführt: bei 40,6 % in der Gruppe der Geburtskliniken der höchsten Versorgungsstufe (Level 3) verglichen mit 46,4 %, 45,5 % und 46,0 % in den Geburtskliniken mit einer niedrigeren Versorgungsstufe (Level 2b, 2a und 1). Aufgrund der aus den relativen Häufigkeiten berechneten absoluten Anzahlen und Abweichungen in der Anzahl der Frauen mit vollständig vorliegenden Informationen wurden in der Auswertung der Versorgungsstufe abweichende Anzahlen von 925 von 6447 Mehrgebärenden mit vaginaler Geburt berechnet, bei denen eine Episiotomie durchgeführt wurde. Von diesen Mehrgebärenden mit vaginaler Geburt wurde bei 12,4 % in der Gruppe der Geburtskliniken der höchsten Versorgungsstufe verglichen mit 15,1 %, 15,4 % und 14,3 % in den Geburtskliniken mit einer niedrigeren Versorgungsstufe (Level 2b, 2a und 1) eine Episiotomie durchgeführt. Die Effektschätzungen nehmen für alle ausgewerteten Subgruppen mit wachsendem Abstand zwischen den Expositionen zu, da mit abnehmender Größe der Geburtskliniken und Versorgungsstufen die Häufigkeit einer Episiotomie konsistent zunimmt.

Eine Metaanalyse zum Einfluss der Erfahrung / Expertise auf die Anzahl der Episiotomien war nicht möglich, da beide Studien auf Daten der französischen Perinatalerhebung 2010 basieren und verschiedene Einschlusskriterien definieren, sodass Coulm 2012 eine Subgruppe der Frauen aus Chuilon 2016 bildet.

Der primär interessierende Vergleich zwischen Geburtskliniken mit einer größeren Geburtenzahl (≥ 3000 Geburten pro Jahr) und einer geringeren Geburtenzahl (< 1000 Geburten pro Jahr) konnte für Frauen mit einer Spontangeburt keinen signifikanten Zusammenhang nachweisen (OR: 0,80; 95 %-KI: [0,57; 1,12]) (siehe Tabelle 57 im Anhang). Beim Vergleich von Kliniken mit einer sehr hohen Geburtenanzahl (> 4000 Geburten pro Jahr), welche in Deutschland sehr selten ist, konnte für Frauen mit einer vaginalen Geburt (ohne Unterscheidung

zwischen einer Spontangeburt und einer vaginal-operativen Geburt) ein statistisch signifikanter Zusammenhang zwischen der Geburtenzahl pro Jahr der Geburtskliniken und der Anzahl von Episiotomien mit einer geringeren Anzahl von Episiotomien in sehr großen Geburtskliniken nachgewiesen werden (RR: 0,71; 95 %-KI: [0,53; 0,96], siehe Abbildung 4 im Anhang).

Beim Vergleich der höchsten und niedrigsten Versorgungsstufe konnte weder für Frauen mit vaginaler Geburt (RR: 1,11; 95 %-KI: [0,91; 1,36]) noch für Frauen mit Spontangeburt (OR: 0,96; 95 %-KI: [0,70; 1,31]) ein signifikanter Unterschied nachgewiesen werden (siehe Abbildung 5 im Anhang).

7.4.2 Wehenmittel

Informationen zum Einfluss der Anzahl der Geburten der Geburtskliniken pro Jahr auf die Häufigkeit von Wehenmittelgaben und zum Einfluss der Versorgungsstufe auf die Häufigkeit von Wehenmittelgaben werden in 1 Studie (Belghiti 2013) berichtet (siehe Tabelle 59 und Tabelle 60 im Anhang).

Insgesamt erhielten 4063 (71,7 %) der ausgewerteten 5668 Frauen aus der französischen Perinatalerhebung mit spontan eingesetzter Wehentätigkeit und PDA Wehenmittel während der Geburt: 62,0 % in der Gruppe der sehr großen Geburtskliniken mit ≥ 4000 Geburten pro Jahr verglichen mit 68,0 %, 71,0 %, 73,0 % und 76,0 % in kleineren Geburtskliniken mit 3000–3999, 2000–2999, 1000–1999 und < 1000 Geburten pro Jahr. Abweichungen in der Anzahl der Frauen mit vollständig vorliegenden Informationen und Nachberechnungen der absoluten Anzahlen aus gerundeten berichteten Angaben zur relativen Häufigkeit bedingen abweichende Anzahlen in der Auswertung zum Einfluss des Versorgungsstatus, nach welcher 4029 (71,2 %) der ausgewerteten 5660 Frauen mit Informationen zum Versorgungsstatus Wehenmittel erhielten. In der Gruppe der Geburtskliniken der höchsten Versorgungsstufe (Level 3) erhielten 66,0 % der Gebärenden ein Wehenmittel während der Geburt, verglichen mit 69,0 %, 73,0 % und 75,0 % der Gebärenden in den Geburtskliniken einer niedrigeren Versorgungsstufe (Level 2b, 2a und 1). Die Effektschätzungen steigen deutlich mit wachsendem Abstand zwischen den Expositionen, da mit abnehmender Größe der Geburtskliniken und Versorgungsstufe das Risiko einer Wehenmittelgabe konsistent zunimmt.

In Belghiti 2013 konnten statistisch signifikante Unterschiede in Abhängigkeit von der Größe der Geburtsklinik und der Versorgungsstufe nachgewiesen werden. Der primär interessierende Vergleich zeigt, dass Gebärende in der Gruppe der sehr großen Geburtskliniken mit ≥ 4000 Geburten im Jahr im Vergleich zu Gebärenden in kleineren Geburtskliniken mit < 1000 Geburten im Jahr signifikant seltener Wehenmittel verabreicht bekamen (OR: 0,67; 95 %-KI: [0,56; 0,83]). Alle in der Studie berichteten Ergebnisse basieren auf Vergleichen zu sehr großen Geburtskliniken mit Geburtenzahlen ≥ 4000 pro Jahr, welche in Deutschland selten sind. Diese Vergleiche werden für den Endpunkt Wehenmittel dennoch herangezogen, da die Unterschiede zwischen den verschiedenen Klinikgrößen hinsichtlich der Wehenmittelgabe deutlich sind und als Ursache-Wirkungs-Zusammenhang interpretiert werden können. Die Häufigkeit von Wehenmittelgaben steigt von den kleinen Geburtskliniken mit < 1000 Geburten

pro Jahr zu den sehr großen Geburtskliniken um 12 Prozentpunkte. Daher wurde der Unterschied von 8 Prozentpunkten zwischen Geburtskliniken mit 3000–3999 Geburten und < 1000 Geburten pro Jahr ebenfalls als relevant angesehen (siehe Tabelle 59 im Anhang).

Analog wurden in den Geburtskliniken der höchsten Versorgungsstufe im Vergleich zu Geburtskliniken mit niedrigerer Versorgungsstufe signifikant seltener Wehenmittel verabreicht (OR: 0,63; 95 %-KI: [0,51; 0,77]).

8 Evidenzprofil für die kritischen und wichtigen Endpunkte

Im Folgenden wird das Evidenzprofil für die kritischen Endpunkte für die Vergleiche zwischen größerer und geringerer Erfahrung und Expertise beschrieben. Da sowohl randomisierte kontrollierte Studien zum Einfluss von Weiterbildungsmaßnahmen auf die Ausprägungen von Erfahrung und Expertise der ärztlichen Geburtshelferinnen oder Geburtshelfer als auch Beobachtungsstudien auf der Basis von Geburtsregistern und Perinataldaten zum Vergleich verschiedener Ausprägungen systemimmanenter Faktoren von Erfahrung und Expertise eingeschlossen wurden, werden die Ergebnisse in 3 getrennten Evidenzprofilen dargestellt.

Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Frage: Größere Erfahrung / Expertise verglichen mit geringerer Erfahrung / Expertise des geburtshilflichen Fachpersonals ab der aktiven Eröffnungsphase											
Setting: klinische Betreuung unter der Geburt											
Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	relatives Risiko [KI]	Absolut [KI]	
Maternale Mortalität											
2	RCT	nicht schwer-wiegend	nicht schwer-wiegend	schwer-wiegend ^a	sehr schwer-wiegend ^b	–	3 / 47 630 (6,3 pro 100 000)	1 / 57 129 (1,8 pro 100 000)	aOR: 2,08; 95 %-KI: [0,05; 81,6] ^c	2 mehr pro 100 000 [von 2 weniger bis 145 mehr] ^d	⊕○○○ SEHR NIEDRIG
Geburtsmodus											
Vaginale Geburt^e											⊕⊕⊕○ MODERAT
2	RCT	schwer-wiegend ^f	nicht schwer-wiegend	nicht schwer-wiegend	nicht schwer-wiegend	–	18 605 / 21 116 (88,1 %)	22 193 / 25 284 (87,8 %)	OR: 1,07 97,5 %-KI: [1,00; 1,14]	1 mehr pro 100 [von 0 bis 3 mehr] ^g	
Kaiserschnitt^h											
2	RCT	schwer-wiegend ^f	nicht schwer-wiegend	nicht schwer-wiegend	nicht schwer-wiegend	–	2511 / 21 116 (11,9 %)	3091 / 25 284 (12,2 %)	OR 0,89; 97,5 %-KI: [0,78; 1,01]	2 weniger pro 100 [von 0 mehr bis 5 weniger] ^g	

(Fortsetzung)

Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	relatives Risiko [KI]	Absolut [KI]	
Höhergradige Dammrisse (III°, IV°)											
1	RCT	schwer-wiegend ⁱ	nicht relevant	nicht schwer-wiegend	schwer-wiegend ^j	–	1241 / 20 612 (6,0 %)	1594 / 24 874 (6,4 %)	aOR: 1,07 [0,90; 1,26]	1 mehr pro 1000 aller Geburten [von 2 weniger bis 5 mehr] ^g	⊕⊕○○ NIEDRIG
										1 mehr pro 1000 vaginale Geburten [von 1 weniger bis 3 mehr] ^g	
Hoher mütterlicher Blutverlust ≥ 1000 ml – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Aufnahme Stillbeziehung bei initialem Stillwunsch – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Teilweise / vollständige Plazentaretention – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Geburtserleben – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Subjektive Gesundheitseinschätzung – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–

(Fortsetzung)

Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	relatives Risiko [KI]	Absolut [KI]	
Neonatale Mortalität											
2	RCT	schwer-wiegend ^f	nicht schwer-wiegend	schwer-wiegend ^a	nicht schwer-wiegend	–	20 / 23 902 (0,08 %)	28 / 29 211 (0,1 %)	aOR: 0,34; 95 %-KI: [0,18; 0,65] ^c	1 weniger pro 10 000 [1 bis 2 weniger] ^g	⊕⊕○○ NIEDRIG
Apgar-Score nach 5 Minuten < 6 – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Schwere metabolische Acidose (pH < 7,0)											
1	RCT	nicht schwer-wiegend	nicht relevant	schwer-wiegend ^a	nicht schwer-wiegend	–	87 / 23 902 (0,36)	90 / 29 211 (0,31)	aOR: 1,13; 95 %-KI: [0,79; 1,62]	0 weniger pro 1000 [von 0 weniger bis 1 mehr] ^g	⊕⊕⊕○ MODERAT
Basendefizit ≥ – 16 mmol/l – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Verlegung auf die Neugeborenen-Intensivstation – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–

(Fortsetzung)

Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	relatives Risiko [KI]	Absolut [KI]	
Episiotomie											
1	RCT	schwer-wiegend ⁱ	nicht relevant	nicht schwer-wiegend	nicht schwer-wiegend	–	2953 / 20 612 (14,3)	3871 / 24 874 (15,6)	aOR: 1,01 95 %-KI: [0,85; 1,21]	0 mehr pro 100 [von 2 weniger bis 3 mehr] ^g	⊕⊕⊕○ MODERAT
Wehenmittel											
1	RCT	schwer-wiegend ⁱ	nicht relevant	nicht schwer-wiegend	nicht schwer-wiegend	–	Wehenmittel während der Geburt				⊕⊕⊕○ MODERAT
							6205 / 20 612 (30,1)	7572 / 24 874 (30,4)	aOR: 1,16 95 %-KI: [1,09; 1,23]	3 mehr pro 100 [von 2 bis 4 mehr] ^g	
							Wehenmittel zur Einleitung der Geburt				
							5501 / 23 484 (23,4)	7872 / 28 781 (27,4)	aOR: 0,82 95 %-KI: [0,76; 0,87]	3 weniger pro 100 [von 2 bis 4 weniger] ^g	
Amniotomie											
1	RCT	schwer-wiegend ⁱ	nicht relevant	schwer-wiegend ^d	nicht schwer-wiegend	–	11 972 / 23 484 (51,0)	14 534 / 28 781 (50,5)	aOR: 1,02 95 %-KI: [0,84; 1,22]	0 mehr pro 100 [von 4 weniger bis 5 mehr] ^d	⊕⊕○○ SEHR NIEDRIG

(Fortsetzung)

Tabelle 17: GRADE-Evidenzprofil zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	relatives Risiko [KI]	Absolut [KI]	
Infektionen											
1	RCT	nicht schwer-wiegend	nicht relevant	schwer-wiegend ^a	schwer-wiegend ^j	–	1986 / 23 484 (8,5)	2337 / 28 781 (8,1)	aOR: 0,95 95 %-KI: [0,75; 1,19]	4 weniger pro 1000 [von 19 weniger bis 14 mehr] ^d	⊕⊕○○ NIEDRIG
Fortgeschrittene metabolische Acidose (pH 7,09–7,0)											
1	RCT	schwer-wiegend ^f	nicht relevant	schwer-wiegend ^a	nicht schwer-wiegend	–	396 / 23 902 (1,7)	413 / 29 211 (1,4)	aOR: 1,49 95 %-KI: [1,02; 2,17]	1 mehr pro 100 [von 0 bis 2 mehr] ^g	⊕⊕○○ NIEDRIG
<p>a: Ergebnis basiert auf allen Gebärenden, von denen 13 % per geplantem Kaiserschnitt gebären.</p> <p>b: Das 95 %-KI lässt sowohl einen bedeutenden Nutzen als auch einen sehr großen Schaden zu. Es erfolgte eine Abwertung.</p> <p>c: Ergebnis für 1 Studie, in 1 weiteren Studie trat bei 914 Gebärenden kein maternaler oder neonataler Todesfall auf.</p> <p>d: Angabe zum Populationsrisiko basiert auf dem Kontrollrisiko der eingeschlossenen Studie aufgrund fehlender Angaben in der Bundesauswertung.</p> <p>e: Aufgrund der fehlenden Differenzierungsmöglichkeit zwischen vaginalen und vaginal-operativen Geburten in 1 der 2 Studien erfolgt die Darstellung der vaginalen Geburten unter Zusammenfassung der Spontangeburt und vaginal-operativen Geburten.</p> <p>f: Fehlende oder unklare Verblindung der behandelnden oder endpunkterhebenden Personen, Effektschätzer wurden ohne Berücksichtigung der Clusterstruktur berechnet. Es erfolgte eine Abwertung.</p> <p>g: Angabe zum Populationsrisiko basiert auf den Daten aus Bundesauswertung IQTIG-Geburtshilfe [5] und bezieht sich auf alle reifgeborenen Kinder.</p> <p>h: Extrahiert wurde nur die Anzahl der ungeplanten Kaiserschnitte im Geburtsverlauf.</p> <p>i: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen. Es erfolgte eine Abwertung.</p> <p>j: Das 95 %-KI lässt sowohl einen relevanten Nutzen als auch Schaden zu. Es erfolgte eine Abwertung.</p> <p>aOR: adjustiertes Odds Ratio; GRADE: Grading of Recommendations, Assessment, Development, and Evaluation (Bewertung von Empfehlungen, Einschätzung, Entwicklung und Auswertung); KI: Konfidenzintervall; OR: Odds Ratio</p>											

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Frage: Größere Erfahrung / Expertise verglichen mit geringerer Erfahrung / Expertise des geburtshilflichen Fachpersonals ab der aktiven Eröffnungsphase											
Setting: klinische Betreuung unter der Geburt											
Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Maternale Mortalität – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Geburtsmodus											
Spontangeburt – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Vaginal-operative Geburt											⊕○○○ SEHR NIEDRIG
1	Beob-achtungs-studie	schwer-wiegend ^a	nicht relevant	nicht schwer-wiegend	nicht schwer-wiegend	Dosis-Wir-kungs-Bezie-hung ^b	Erst- und Mehrgebärende: ≥ 3000 vs. < 1000				
							k. A. (16,5 %)	k. A. (12,7 %)	aOR: 1,47; 98,3 %-KI: [1,03; 2,09]	5 mehr pro 100 [von 0 bis 11 mehr] ^c	

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Vaginale Geburt											⊕○○○ SEHR NIEDRIG
1	Beob-achtungs studie	schwer-wiegend ^{a, d}	nicht relevant	für Erstge-bärende: nicht schwer-wiegend	nicht schwer-wiegend ^f	–	Erstgebärende: > 3000 vs. 1100–3000				
							36 950 / 41 689 (88,6 %)	24 554 / 27 970 (87,8 %);	aOR: 1,22; 97,5 %-KI: [1,14; 1,30]	2 mehr pro 100 [von 1 bis 3 mehr] ^c	
				für Mehrge-bärende: schwer-wiegend ^e			Erstgebärende: 1100–3000 vs. < 1100				
							24 554 / 27 970 (87,8 %)	3800 / 4343 (87,5 %)	aOR: 1,02; 97,5 %-KI: [0,91; 1,15]	0 weniger pro 100 [von 1 weniger bis 1 mehr] ^c	
Kaiserschnitt											
1	Beob-achtungs studie	schwer-wiegend ^a	nicht relevant	für Erstge-bärende: nicht schwer-wiegend	nicht schwer-wiegend ^f	–	Erstgebärende: > 3000 vs. 1100–3000				
							4739 / 41 689 (11,4 %)	3416 / 27 970 (12,2 %)	aOR: 0,82; 98,3 %-KI: [0,77; 0,88]	2 weniger pro 100 [von 1 bis 3 weniger] ^c	
				für Mehrge-bärende: schwer-wiegend ^e			Erstgebärende: 1100–3000 vs. < 1100				
							3416 / 27 970 (12,2 %)	543 / 4343 (12,5 %)	aOR: 0,98; 98,3 %-KI: [0,86; 1,12]	0 weniger pro 100 [von 2 weniger bis 1 mehr] ^c	

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien- design	Studien- limitationen	Inkon- sistenz	Indirekt- heit	Fehlende Genauig- keit	Andere Fak- toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Höhergradige Dammrisse (III°, IV°) – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Hoher mütterlicher Blutverlust ≥ 1000 ml – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Aufnahme Stillbeziehung bei initialem Stillwunsch – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Teilweise / vollständige Plazentaretention – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Geburtserleben – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Subjektive Gesundheitseinschätzung – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
neonatale Mortalität											
2	Beob-achtungs studien	sehr schwer-wiegend ^g	nicht zu-treffend ^h	schwer-wiegend ⁱ	sehr schwer-wiegend ^j	–	1000–2499 vs. < 500				⊕○○○ SEHR NIEDRIG
							k. A. / 638 183	k. A. / 31 745	aOR: 1,19; 95 %-KI: [0,90; 1,59]	0 mehr pro 10 000 [von 0 weniger bis 1 mehr] ^k	
							Universitätskrankenhaus vs. ≤ 999 (frühe neonatale Todesfälle 1.–7. Lebenstag)				
							14 / 92 139 (0,02 %)	12 / 39 385 (0,03 %)	aOR: 0,47; 95 %-KI: [0,22; 1,03]	2 weniger pro 10 000 [von 2 weniger bis 0 mehr] ^c	
Universitätskrankenhaus vs. ≤ 999 (intrapartale Todesfälle)											
							219 / 92 139 (0,24 %)	108 / 39 385 (0,27 %)	aOR: 0,84; 95 %-KI: [0,66; 1,05]	4 weniger pro 10 000 [von 9 weniger bis 1 mehr] ^c	
Apgar-Score nach 5 Minuten < 6 – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Risiko [KI]	Absolut [KI]	
Schwere metabolische Acidose (pH < 7,0)											
1	Beob-achtungs studie	sehr schwer-wiegend ^g	nicht relevant	nicht schwer-wiegend	nicht schwer-wiegend	–	Universitätskrankenhaus vs. ≤ 999				
							361 / 92 139 (0,39 %)	99 / 39 385 (0,25 %)	aOR: 1,43; 95 %-KI: [1,14; 1,79]	11 mehr pro 10 000 [von 3 bis 20 mehr] ^c	⊕○○○ SEHR NIEDRIG
Basendefizit ≥ – 16 mmol/l – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Verlegung auf die Neugeborenen-Intensivstation – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Episiotomie											
1	Beob-achtungs studie	schwer-wiegend ^a	nicht relevant	schwer-wiegend ^l	schwer-wiegend ^m	–	≥ 3000 vs. < 1000				
							209 / 1227 (17,0 %)	269 / 1354 (19,9 %)	aOR: 0,80; 95 %-KI: [0,57; 1,12]	3 weniger pro 100 [von 7 weniger bis 2 mehr] ^c	⊕○○○ SEHR NIEDRIG

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Risiko [KI]	Absolut [KI]	
Wehenmittel											
1	Beob-achtungs studie	schwer-wiegend ^a	nicht relevant	schwer-wiegend ^a	nicht schwer-wiegend	–	≥ 4000 vs. < 1000		aOR: 0,67; 95 %-KI: [0,56; 0,83]	8 weniger pro 100 [von 4 bis 12 weniger] ^b	⊕○○○ SEHR NIEDRIG
							159 / 256 (62,0 %)	695 / 914 (76,0 %)		7 weniger pro 100 [von 3 bis 10 weniger] ^o	

(Fortsetzung)

Tabelle 18: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr) (Fortsetzung)

- a: Fehlende Verblindung von Gebärenden und behandelnden Personen, unklare ergebnisunabhängige Berichterstattung. Es erfolgte eine Abwertung.
- b: Dosis-Wirkungs-Beziehung beschreibt hier den Zusammenhang zwischen Klinikgröße und dem Auftreten des Endpunkts. Mit abnehmender Geburtenzahl sinkt die Häufigkeit der Ausprägung des Endpunkts in jeweils 1 Studie.
- c: Das Populationsrisiko beruht auf dem Risiko aus der Referenzgruppe.
- d: Unterschiedliche Erfassung bestimmter Risikofaktoren ohne präzise Definition kann nicht ausgeschlossen werden, da diese erst nach der Geburt erfasst wurden.
- e: In der Studie liegen ausschließlich Ergebnisse für Erstgebärende vor.
- f: Es erfolgte keine Abwertung, obwohl das KI des absoluten Effekts teilweise sowohl einen relevanten Nutzen als auch Schaden einschließt, dieser Unterschied aufgrund der absoluten Effekte aber als nicht klinisch relevant eingestuft wurde.
- g: unklare adäquate Confounderadjustierung bei fehlenden Informationen zu Baseline-Charakteristika in Expositions- und Referenzgruppe, fehlende Verblindung von Gebärenden und behandelnden Personen, unklare ergebnisunabhängige Berichterstattung. Der Ausschluss von 324 (0,12 %) der Neugeborenen aufgrund fehlender Werte kann den Effektschätzer für dieses seltene Ereignis beeinflussen. Es erfolgte eine Abwertung.
- h: Die Ergebnisse wurden nicht zusammengefasst, da unterschiedliche Zeiträume betrachtet wurden.
- i: In 1 Studie wird das Ereignis für alle lebend geborenen Einlinge berichtet. Der Einschluss von vor der vollendeten 37. SSW geborenen Neugeborenen kann den Effektschätzer beeinflussen.
- j: Die Punktschätzer und 95 %-KI umfassen sowohl einen bedeutenden Nutzen als auch einen sehr großen Schaden. Es erfolgte eine Abwertung.
- k: Angabe zum Populationsrisiko basiert auf Daten aus Bundesauswertung IQTIG-Geburtshilfe [5] und bezieht sich auf alle reifgeborenen Kinder.
- l: In der Studie liegen ausschließlich Ergebnisse für Gebärende mit einer Spontangeburt vor. Informationen zu Gebärenden mit einer vaginal-operativen Geburt fehlen.
- m: Die Punktschätzer und 95 %-KI umfassen sowohl einen relevanten Nutzen als auch Schaden. Es erfolgte eine Abwertung.
- n: In der Studie liegen ausschließlich Ergebnisse für Gebärende mit einer laut Autorengruppe für Frankreich typischen Standardbetreuung (spontan eingesetzte Wehentätigkeit und PDA) und für einen Vergleich zu sehr großen Kliniken vor, welche in Deutschland sehr selten sind. Es erfolgte eine Abwertung.
- o: Angabe zum Populationsrisiko basiert auf Daten aus Bundesauswertung IQTIG-Geburtshilfe [5], da das Risiko aus der Referenzgruppe relevant von dem deutschen Versorgungskontext abweicht, und bezieht sich auf Wehenmittelgabe während der Geburt. Es erfolgte eine Abwertung.
- aOR: adjustiertes Odds Ratio; GRADE: Grading of Recommendations, Assessment, Development, and Evaluation (Bewertung von Empfehlungen, Einschätzung, Entwicklung und Auswertung); k. A. keine Angabe; KI: Konfidenzintervall; PDA: Periduralanästhesie; RD: Risikodifferenz; SSW: Schwangerschaftswoche; vs.: versus

Tabelle 19: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe)

Frage: Größere Erfahrung / Expertise verglichen mit geringerer Erfahrung / Expertise des geburtshilflichen Fachpersonals ab der aktiven Eröffnungsphase											
Setting: klinische Betreuung unter der Geburt											
Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Maternale Mortalität – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Geburtsmodus											
Spontangeburt											⊕○○○ SEHR NIEDRIG
Vaginale-operative Geburt^a											
1	Beobach-tungs-studie	sehr schwer-wiegend ^a	nicht relevant	nicht schwer-wiegend	schwer-wiegend ^b	keine	k. A. (15,0 %)	k. A. (15,9 %)	aOR: 1,30; 98,3 %-KI: [0,95; 1,77]	4 mehr pro 100 [von 1 weniger bis 9 mehr] ^c	
Kaiserschnitt – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Höhergradige Dammrise (III°, IV°) – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Hoher mütterlicher Blutverlust ≥ 1000 ml – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Aufnahme Stillbeziehung bei initialem Stillwunsch – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–

(Fortsetzung)

Tabelle 19: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe) (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Teilweise / vollständige Plazentaretention – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Geburtserleben – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
subjektive Gesundheitseinschätzung – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
neonatale Mortalität – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Apgar-Score nach 5 Minuten < 6 – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
schwere metabolische Acidose (pH < 7,0) – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Basendefizit ≥ – 16 mmol/l – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Verlegung auf die Neugeborenen-Intensivstation – keine Angaben											
–	–	–	–	–	–	–	–	–	–	–	–
Episiotomie											
1 ^d	Beob-achtungs-studie	schwer-wiegend ^a	nicht relevant	nicht schwer-wiegend	schwer-wiegend ^e	Dosis-Wir-kungs-Bezie-hung ^f	600 / 2463 (24,4 %)	928 / 3358 (27,6 %)	aRR: 1,11; 95 %-KI: [0,91; 1,36]	3 mehr pro 100 [von 2 weniger bis 10 mehr] ^c	⊕○○○ SEHR NIEDRIG

(Fortsetzung)

Tabelle 19: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe) Tabelle 19: GRADE-Evidenzprofil zur Erfahrung / Expertise durch systemimmanente Faktoren (Fortsetzung)

Bewertung der Qualität der Evidenz							Anzahl der Ereignisse / Gebärende		Wirkung (Effekt)		Qualität der Evidenz
Anzahl der Studien	Studien-design	Studien-limitationen	Inkon-sistenz	Indirekt-heit	Fehlende Genauig-keit	Andere Fak-toren	größere Expertise	kleinere Expertise	Odds Ratio [KI]	Absolut [KI]	
Wehenmittel											
1	Beob-achtungs studie	schwer-wiegend ^a	nicht relevant	schwer-wiegend ^b	nicht schwer-wiegend	Dosis-Wir-kungs-Bezie-hung ^f	832 / 1261 (66,0 %)	1244 / 1659 (75,0 %)	aOR: 0,63; 95 %-KI: [0,51; 0,77]	10 weniger pro 100 [von 15 bis 5 weniger] ^c	⊕○○○ SEHR NIEDRIG
										8 weniger pro 100 [von 11 bis 5 weniger] ^h	
<p>a: unklare ergebnisunabhängige Berichterstattung. Es erfolgte eine Abwertung.</p> <p>b: Das 98,3 %-KI schließt sowohl einen relevanten Nutzen als auch einen Schaden ein. Es erfolgte eine Abwertung.</p> <p>c: Angaben basieren auf dem Populationsrisiko der Referenzgruppe, Berechnung basiert für die neonatale Mortalität auf den Ergebnissen aus 1 Studie, da in der 2. Studie bei insgesamt 914 Neugeborenen erwartungsgemäß kein Ereignis auftrat.</p> <p>d: 1 weitere Studie berichtet Ergebnisse für Frauen mit einer Spontangeburt, einer Subgruppe der vaginal gebärenden Frauen. Es wird das Ergebnis für Frauen mit einer vaginalen Geburt dargestellt, da diese die Zielpopulation besser abbilden.</p> <p>e: Das 95 %-KI schließt sowohl einen relevanten Nutzen als auch einen Schaden ein. Es erfolgte eine Abwertung.</p> <p>f: Dosis-Wirkungs-Beziehung beschreibt hier den Zusammenhang zwischen Versorgungsstufe und dem Auftreten des Endpunkts. Mit zunehmender Versorgungsstufe sinkt die Häufigkeit der Ausprägung des Endpunkts in jeweils 1 Studie.</p> <p>g: In der Studie liegen ausschließlich Ergebnisse für Gebärende mit einer laut Autorengruppe für Frankreich typischen Standardbetreuung (spontan eingesetzte Wehentätigkeit und PDA) vor. Es erfolgte eine Abwertung.</p> <p>h: Angabe zum Populationsrisiko basiert auf Daten aus Bundesauswertung IQTIG-Geburtshilfe [5], da das Risiko aus der Referenzgruppe relevant von dem deutschen Versorgungskontext abweicht, und bezieht sich auf Wehenmittelgabe während der Geburt. Es erfolgte eine Abwertung.</p> <p>aOR: adjustiertes Odds Ratio; aRR: adjustiertes relatives Risiko; GRADE: Grading of Recommendations, Assessment, Development, and Evaluation (Bewertung von Empfehlungen, Einschätzung, Entwicklung und Auswertung); k. A.: keine Angaben; KI: Konfidenzintervall; PDA: Periduralanästhesie;</p>											

8.1 Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Im Folgenden wird die vorliegende Evidenz zu den 5 berichteten kritischen und 5 wichtigen Endpunkten der 2 eingeschlossenen clusterrandomisierten Studie zusammenfassend dargestellt.

Die Angaben zum absoluten Effekt basieren für die neonatale Mortalität, den Geburtsmodus, das Auftreten höhergradiger Dammrisse (III°, IV°), die schwere metabolische Acidose (pH < 7,0), Episiotomie, Wehenmittel und die fortgeschrittene metabolische Acidose (pH 7,09–7,0) auf dem Populationsrisiko aus den Angaben der Bundesauswertung zur stationären Qualitätssicherung zum Erfassungsjahr 2017 des Instituts für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) [5]. Für die Endpunkte maternale Mortalität, Amniotomie sowie Infektionen basiert aufgrund fehlender Angaben in der IQTIG-Auswertung das Populationsrisiko auf den Kontrollgruppenrisiken der eingeschlossenen Studie mit berichteten Ergebnissen.

Die Evidenzqualität der RCTs wurde aufgrund schwerwiegender Studienlimitationen, Indirektheit und (sehr) schwerwiegender fehlender Genauigkeit herabgestuft.

Schwerwiegende Studienlimitationen, welche zur Herabstufung um 1 Kategorie führten, umfassten

- eine unklare Verblindung der Teilnehmerinnen und eine fehlende Verblindung der behandelnden und endpunkterhebenden Personen
- eine Berechnung von Effektschätzern ohne Berücksichtigung der Clusterstruktur

Schwerwiegende Indirektheit führte zur Herabstufung um 1 Kategorie, wenn die gebärenden Frauen von der relevanten Population abweichen und die Ergebnisse für alle gebärenden Frauen, einschließlich der Frauen mit geplantem Kaiserschnitt, berichtet wurden.

Schwerwiegende fehlende Genauigkeit führte zur Abwertung der Evidenz für einen Endpunkt, wenn die Ergebnisse zu unpräzise waren, um eine Entscheidung für oder gegen eine Empfehlung stützen zu können. Primär basierte eine Abwertung um 1 Kategorie auf einem KI, das sowohl einen relevanten Nutzen als auch einen Schaden einschließt. Es erfolgte eine Abwertung um 2 Kategorien, wenn das angegebene KI sowohl einen bedeutenden Nutzen als auch einen sehr großen Schaden zulässt.

Es konnte für keinen der untersuchten Endpunkte ein Hinweis auf einen Publikationsbias festgestellt werden, sodass **keine Abwertung wegen des Verdachts auf Publikationsbias** erfolgte.

8.1.1 Maternale Mortalität

Die Evidenz stammt aus 2 Studien (Chaillet 2015, Lomas 1991). Insgesamt verstarben 3 von 47 630 Gebärenden während der Prä- und Postinterventionsperiode in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, verstarb 1 von 57 129 Gebärenden. Die Ergebnisse dieser Studie gingen in die Berechnung der Effekte ein. Diese basieren auf den Daten aller Gebärender, von denen 12,8 % per geplantem Kaiserschnitt entbanden. In 1 weiteren Studie verstarb keine der 914 Gebärenden.

Die Qualität der Evidenz wurde für den Endpunkt maternale Mortalität als sehr niedrig bewertet, da aufgrund schwerwiegender Indirektheit und sehr schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 17). Diese Herabstufung basierte auf Ungenauigkeiten der Effektschätzung. Das 95 %-KI des relativen Risikos (OR: 2,08; 95 %-KI: [0,05; 81,6]) schließt sowohl einen bedeutenden Nutzen als auch einen sehr großen Schaden ein.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach verstarben 2017 in Deutschland bezogen auf alle Geburten 2 von 100 000 Frauen.

Eine Veränderung in der maternalen Mortalität konnte nicht nachgewiesen werden (OR: 2,08; 95 %-KI: [0,05; 81,6]).

8.1.2 Geburtsmodus

Insgesamt traten in den Studien 18 605 vaginale Geburten und 2511 ungeplante Kaiserschnitte im Geburtsverlauf bei insgesamt 21 116 geplanten vaginalen Geburten in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurde, auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, gebaren 22 193 der insgesamt 25 284 Frauen vaginal und 3091 Frauen wiesen einen ungeplanten Kaiserschnitt auf. Eine Differenzierungsmöglichkeit zwischen Spontangeburt und vaginal-operativer Geburt ist aufgrund der Angaben in der Publikation von 1 Studie nicht möglich. Für den Kaiserschnitt werden die Ergebnisse für den ungeplanten Kaiserschnitt im Geburtsverlauf und nicht für den geplanten Kaiserschnitt bewertet, da die Entscheidung für einen geplanten Kaiserschnitt per Definition in der Regel bereits vor Beginn der aktiven EP getroffen wird und somit Frauen mit einem geplanten Kaiserschnitt nicht der Zielpopulation dieses Evidenzberichts entsprechen.

Die Qualität der Evidenz wurde für den Endpunkt Geburtsmodus zusammenfassend als moderat bewertet, da aufgrund schwerwiegender Studienlimitationen herabgestuft wurde (siehe

Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasst eine unklare Verblindung der teilnehmenden und eine fehlende Verblindung der behandelnden und endpunkterhebenden Personen sowie eine fehlende Berücksichtigung der Clusterstruktur für die vaginale Geburt.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach erfolgten 2017 in Deutschland 69 pro 100 Geburten vaginal und 31 pro 100 per Kaiserschnitt.

Für die 2 berichteten Ausprägungen des Endpunkts Geburtsmodus konnte zwar für die Anzahl von vaginalen Geburten (Spontangeburt und vaginal-operative Geburten) ein signifikanter Einfluss (RR: 1,07; 97,5 %-KI: [1,00; 1,14]), aber nicht auf die Anzahl ungeplanter Kaiserschnitte (RR: 0,89; 97,5 %-KI: [0,78; 1,01]) nachgewiesen werden. Demnach traten nach Einsatz der Weiterbildungsmaßnahme 1 (von 0 bis 3) mehr vaginale Geburten pro 100 Geburten auf.

8.1.3 Höhergradige Dammrisse (III°, IV°)

Die Evidenz stammt aus 1 Studie (Chaillet 2015). Insgesamt trat bei 1241 von 20 612 Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, ein höhergradiger Dammriss auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, trat bei 1594 von 24 874 Gebärenden ein höhergradiger Dammriss auf.

Die Qualität der Evidenz wurde für den Endpunkt höhergradige Dammrisse (III°, IV°) als niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasst eine unklare Verblindung der teilnehmenden und eine fehlende Verblindung der behandelnden und endpunkterhebenden Personen. Die weitere Herabstufung basierte auf Ungenauigkeiten der Effektschätzung. Das 95 %-KI des relativen Effektes (OR: 1,07; 95 %-KI: [0,90; 1,26]) schließt sowohl einen relevanten Nutzen als auch einen Schaden ein.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach trat 2017 in Deutschland bezogen auf alle Geburten bei 13 pro 1000 Frauen und bezogen auf vaginale Geburten bei 19 pro 1000 Frauen ein höhergradiger Dammriss auf.

Eine Veränderung im Auftreten höhergradiger Dammrisse (III°, IV°) konnte nicht nachgewiesen werden (OR: 1,07; 95 %-KI: [0,90; 1,26]).

8.1.4 Neonatale Mortalität

Die Evidenz stammt aus 2 Studien (Chaillet 2015, Lomas 1991). Insgesamt verstarben 20 von 23 902 Neugeborenen in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, verstarben 28 von 29 211 Neugeborenen. Die Ergebnisse dieser Studie gingen in die Berechnung der Effekte ein. Diese basieren auf den Daten aller Gebärender, von denen 13,0 % per geplantem Kaiserschnitt entbanden. In 1 weiteren Studie verstarb keines der 914 Neugeborenen.

Die Qualität der Evidenz wurde für den Endpunkt neonatale Mortalität zusammenfassend als niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen ist durch die fehlende Berücksichtigung der Clusterstruktur bedingt.

Die Berechnungen des absoluten Effektes basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach verstarben 2017 in Deutschland bezogen auf alle Geburten 2 von 10 000 reifgeborenen Einlingen.

Es konnte für die neonatale Mortalität für eine größere Erfahrung / Expertise beschrieben durch eine Weiterbildungsmaßnahme ein signifikanter Unterschied im Vergleich zur Vergleichsintervention ohne Weiterbildungsmaßnahmen nachgewiesen werden (OR: 0,34; 95 %-KI: [0,18; 0,65]). Demnach verstarben nach Einsatz der Weiterbildungsmaßnahme 1 (0 bis 3) weniger Neugeborene pro 10 000 Geburten.

8.1.5 Schwere metabolische Acidose (pH < 7,0)

Die Evidenz stammt aus 1 Studie (Chaillet 2015). Insgesamt trat bei 87 von 23 902 Neugeborenen in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine schwere metabolische Acidose auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welchen keine Weiterbildungsmaßnahmen geplant waren, trat bei 90 von 29 211 Neugeborenen eine schwere metabolische Acidose auf. Diese Ergebnisse basieren auf den Daten aller Gebärender, von denen 13,0 % per geplantem Kaiserschnitt entbanden.

Die Qualität der Evidenz wurde für den Endpunkt schwere metabolische Acidose (pH < 7,0) als moderat bewertet, da aufgrund schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 17). Aufgrund der hohen Präzision des absoluten Effektes wird davon ausgegangen, dass die Intervention keinen Einfluss auf das Auftreten einer schweren metabolischen Acidose hat.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach trat 2017 in Deutschland bezogen auf alle lebendgeborenen reifen Einlinge bei 2 pro 1000 Neugeborenen eine schwere metabolische Acidose auf.

Eine Veränderung im Auftreten einer schweren metabolischen Acidose konnte nicht nachgewiesen werden (OR: 1,13; 95 %-KI: [0,79; 1,62]).

8.1.6 Episiotomie

Die Evidenz stammt aus 1 Studie (Chaillet 2015). Insgesamt trat bei 2953 von 20 612 Gebärenden mit angestrebter vaginaler Geburt in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine Episiotomie auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, trat bei 3871 von 24 874 Gebärenden eine Episiotomie auf.

Die Qualität der Evidenz wurde für den Endpunkt Episiotomie als moderat bewertet, da aufgrund schwerwiegender Studienlimitationen herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasst eine unklare Verblindung der teilnehmenden und eine fehlende Verblindung der behandelnden und endpunkterhebenden Personen. Aufgrund der hohen Präzision des absoluten Effektes wird davon ausgegangen, dass die Intervention keinen Einfluss auf das Auftreten von Episiotomien hat.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach trat 2017 in Deutschland bezogen auf alle Geburten bei 19 pro 100 Frauen eine Episiotomie auf.

Eine Veränderung im Auftreten von Episiotomien konnte nicht nachgewiesen werden (OR: 1,01; 95 %-KI: [0,85; 1,21]).

8.1.7 Wehenmittel

Die Evidenz stammt aus 1 Studie (Chaillet 2015), in welcher für alle Gebärenden Ergebnisse zur Einleitung der Geburt und für Gebärende mit angestrebter vaginaler Geburt Ergebnisse zur Gabe von Wehenmitteln während der Geburt berichtet werden.

Von den Gebärenden mit angestrebter vaginaler Geburt erhielten in der Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme 6205 von 20 612 und in der Gruppe ohne geplante Weiterbildungsmaßnahmen 7572 von 24 874 Gebärenden Wehenmittel.

Insgesamt erhielten 5501 von 23 484 Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine Wehenmittelgabe zur

Einleitung der Geburt. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welchen keine Weiterbildungsmaßnahmen geplant waren, erhielten 7872 von 28 781 Gebärenden Wehenmittel zur Geburtseinleitung. Diese Ergebnisse basieren auf den Daten aller Gebärender, von denen 13,0 % per geplantem Kaiserschnitt entbanden. Es wurde auf eine Abwertung der Qualität der Evidenz verzichtet, da zusätzlich Informationen zur Wehenmittelgabe unter der Geburt für Frauen mit angestrebter vaginaler Geburt berichtet wurden.

Die Qualität der Evidenz wurde für den Endpunkt Wehenmittelgaben als moderat bewertet, da aufgrund schwerwiegender Studienlimitationen herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasst eine unklare Verblindung der teilnehmenden und eine fehlende Verblindung der behandelnden und endpunkterhebenden Personen.

Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach erhielten 2017 in Deutschland bezogen auf alle Geburten 22 pro 100 Frauen eine Wehenmittelgabe zur Einleitung der Geburt und 26 pro 100 Frauen Wehenmittel unter der Geburt.

Es konnte für die Gabe von Wehenmitteln für eine größere Erfahrung / Expertise beschrieben durch eine Weiterbildungsmaßnahme ein signifikanter Unterschied im Vergleich zur Vergleichsintervention ohne Weiterbildungsmaßnahmen nachgewiesen werden. Danach wurden nach Weiterbildungsmaßnahmen mehr Geburten eingeleitet (OR: 1,16; 95 %-KI: [1,09; 1,23]), aber seltener Wehenmittel unter der Geburt eingesetzt (OR: 0,82; 95 %-KI: [0,76; 0,87]). Demnach erhielten nach Einsatz der Weiterbildungsmaßnahme 3 (2 bis 4) mehr pro 100 Frauen eine Wehenmittelgabe während der Geburt und 3 (2 bis 4) weniger Frauen erhielten Wehenmittelgaben zur Einleitung der Geburt.

8.1.8 Amniotomie

Die Evidenz stammt aus 1 Studie (Chaillet 2015). Insgesamt erhielten 11 972 von 23 484 Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine Amniotomie. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, wiesen 14 534 von 28 781 Gebärende eine Amniotomie auf. Die berichteten Effekte basieren auf den Daten aller Gebärender, von denen 13,0 % per geplanten Kaiserschnitt entbanden.

Die Qualität der Evidenz wurde für den Endpunkt Amniotomie als niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassen eine unklare Verblindung der teilnehmenden und

eine fehlende Verblindung der behandelnden sowie endpunkterhebenden Personen. Die weitere Herabstufung basierte auf Ungenauigkeiten der Effektschätzung.

Aufgrund der hohen Präzision des absoluten Effektes wird davon ausgegangen, dass die Intervention keinen Einfluss auf das Auftreten von Amniotomien hat.

Da in den Angaben des IQTIG [5] keine Informationen zu Amniotomien enthalten sind, basieren die Berechnungen der absoluten Effekte auf dem Kontrollgruppenrisiko der eingeschlossenen Studie. Demnach trat auf alle Geburten bei 50 pro 100 Frauen eine Amniotomie auf.

Eine Veränderung im Auftreten von Amniotomien konnte nicht nachgewiesen werden (OR: 1,02; 95 %-KI: [0,84; 1,22]).

8.1.9 Infektionen

Die Evidenz stammt aus 1 Studie (Chaillet 2015), in welcher Ergebnisse zum Auftreten von Infektionen / Sepsis berichtet werden.

Insgesamt traten bei 1986 von 23 484 Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine Infektion / Sepsis auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, trat bei 2337 von 28 781 Gebärenden eine Infektion / Sepsis auf. Die berichteten Effekte basieren auf den Daten aller Gebärender, von denen 13,0 % per geplanten Kaiserschnitt entbanden.

Die Qualität der Evidenz wurde für den Endpunkt Infektionen als niedrig bewertet, da aufgrund schwerwiegender fehlender Genauigkeit und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 17). Die Herabstufung basierte auf Ungenauigkeiten der Effektschätzung. Das 95 %-KI des relativen Effektes (OR: 0,95; 95 %-KI: [0,75; 1,19]) schließt sowohl einen relevanten Nutzen als auch einen Schaden ein.

Da in den Angaben des IQTIG [5] ausschließlich Informationen zu Sepsis vorliegen und sich das Kontrollgruppenrisiko aus der Studie wesentlich von dem aus den IQTIG-Daten unterscheidet, basieren die Berechnungen der absoluten Effekte auf dem Kontrollgruppenrisiko der eingeschlossenen Studie. Demnach trat bei 8 von 100 Gebärenden eine Infektion / Sepsis auf.

Eine Veränderung im Auftreten von Infektionen konnte nicht nachgewiesen werden (OR: 0,95; 95 %-KI: [0,75; 1,19]).

8.1.10 Fortgeschrittene metabolische Acidose

Die Evidenz stammt aus 1 Studie (Chaillet 2015). Insgesamt trat bei 396 von 23 902 Neugeborenen in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Erfahrung / Expertise, welche einer Interventionsgruppe mit einer geplanten Weiterbildungsmaßnahme zugewiesen wurden, eine fortgeschrittene metabolische Acidose auf. In der Gruppe mit einer Betreuung in einer Geburtsklinik mit kleinerer Erfahrung / Expertise, in welcher keine Weiterbildungsmaßnahmen geplant waren, trat bei 413 von 29 211 Neugeborenen eine fortgeschrittene metabolische Acidose auf. Die berichteten Effekte basieren auf den Daten aller Gebärender, von denen 13,0 % per geplanten Kaiserschnitt entbanden.

Die Qualität der Evidenz wurde für den Endpunkt fortgeschrittene metabolische Acidose (pH 7,09–7,0) als niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 17). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, ist durch die fehlende Berücksichtigung der Clusterstruktur bedingt. Die Berechnungen der absoluten Effekte basieren auf dem Populationsrisiko aus den Angaben des IQTIG [5]. Demnach trat 2017 in Deutschland bezogen auf alle lebendgeborenen reifen Einlinge bei 2 pro 100 Neugeborenen eine fortgeschrittene metabolische Acidose auf.

Es konnte für die fortgeschrittene metabolische Acidose für eine größere Erfahrung / Expertise beschrieben durch eine Weiterbildungsmaßnahme ein signifikanter Unterschied im Vergleich zur Vergleichsintervention ohne Weiterbildungsmaßnahmen nachgewiesen werden (OR: 1,49; 95 %-KI: [1,02; 2,17]). Demnach wiesen nach Einsatz der Weiterbildungsmaßnahme 1 (0 bis 2) mehr Neugeborene pro 100 Geburten eine fortgeschrittene metabolische Acidose auf.

8.1.11 Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte

Es konnten auf der Grundlage von RCTs Effektschätzungen für insgesamt 5 kritische und 5 wichtige Endpunkte ermittelt werden. Die Qualität der Evidenz wurde für 4 Endpunkte als moderat (Geburtsmodus, schwere metabolische Acidose [pH < 7,0], Episiotomie, Wehenmittel), für 5 Endpunkte als niedrig (höhergradige Dammrisse [III°, IV°], neonatale Mortalität, Amniotomie, Infektionen, fortgeschrittene Acidose) und für 1 Endpunkt als sehr niedrig (maternale Mortalität) bewertet. Ergebnisse aus RCTs zu hohem mütterlichen Blutverlust ≥ 1000 ml, zur Aufnahme einer Stillbeziehung bei initialem Stillwunsch, zur teilweisen / vollständigen Plazentaretention, zum Geburtserleben, zur subjektiven Gesundheitseinschätzung, zum Apgar-Score nach 5 Minuten < 6, zum Basendefizit ≥ -16 mmol/l und zur Verlegung auf die Neugeborenen-Intensivstation wurden nicht berichtet.

Zusammenfassend wurde die Qualität der Evidenz als niedrig bewertet, sodass folglich das Gesamtvertrauen in die Effektschätzungen begrenzt ist und der wahre Effekt einer größeren Erfahrung / Expertise beschrieben durch Weiterbildungsmaßnahmen sich von den berechneten Effektschätzungen relevant unterscheiden kann.

8.2 Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Im Folgenden wird die vorliegende Evidenz zu den 3 berichteten kritischen Endpunkten und den 2 berichteten wichtigen Endpunkten der eingeschlossenen Studien zur Geburtenzahl pro Jahr der Geburtsklinik zusammenfassend dargestellt.

Die Angaben zum absoluten Effekt basieren auf dem Risiko der Referenzgruppe mit geringerer Erfahrung / Expertise, da Angaben der Bundesauswertung zur stationären Qualitätssicherung zum Erfassungsjahr 2017 des Instituts für Qualitätssicherung und Transparenz (IQTIG) [5] keine Differenzierungsmöglichkeit zwischen den verschiedenen Geburtenzahlen der Geburtskliniken enthalten. Für den Endpunkt neonatale Mortalität und Wehenmittel wurden zusätzlich für jeweils 1 Studie IQTIG-Angaben als Populationsrisiko herangezogen, da Angaben zum Risiko der Referenzgruppe nicht berichtet werden oder das Populationsrisiko für diesen Endpunkt sich in Deutschland bedeutend von dem des Risikos der Referenzgruppe aus der Studie unterscheidet.

Erfahrung / Expertise wird auf Basis der Geburtenzahlen in den Geburtskliniken beschrieben, wobei Geburtskliniken mit geringeren Geburtenzahlen die Referenzgruppe bilden.

Ausgehend von einer niedrigen Qualität der Evidenz, bedingt durch das Studiendesign der eingeschlossenen Studien, erfolgte eine Herabstufung aufgrund sehr schwerwiegender Studienlimitationen, Inkonsistenz, Indirektheit und schwerwiegender fehlender Genauigkeit.

Schwerwiegende Studienlimitationen mit der Möglichkeit einer Herabstufung der Qualität der Evidenz um 2 Kategorien umfassten

- unklare adäquate Confounderadjustierung bei fehlenden Informationen zu Baseline-Charakteristika in den Expositions- und Referenzgruppen,
- eine fehlende Verblindung der Teilnehmerinnen und der behandelnden Personen,
- eine unklare oder fehlende ergebnisunabhängige Berichterstattung, da Angaben zu geplanten primären und sekundären Endpunkten oder eine Differenzierung diesbezüglich fehlen, keine Studienprotokolle vorliegen und in 1 Studie eine selektive Berichterstattung signifikanter Ergebnisse vorliegt,
- eine möglicherweise unterschiedliche Erfassung von Risikofaktoren ohne standardisierte Definition, welche in die Confounderadjustierung einfließen,
- der Ausschluss von Neugeborenen aufgrund fehlender Daten aus der Analyse und die mögliche Beeinflussung des Ergebnisses für seltene Ereignisse.

Eine Abwertung aufgrund **schwerwiegender Inkonsistenz** erfolgte nicht, da keine metaanalytische Zusammenfassung der Ergebnisse verschiedener Einzelstudien erfolgte. Wenn Ergebnisse auf 1 Studie basieren, wurde die Inkonsistenz als nicht relevant beschrieben.

Eine Abwertung aufgrund **schwerwiegender Indirektheit** mit der Möglichkeit einer Herabstufung der Qualität der Evidenz um 1 bis 2 Kategorien erfolgte, wenn ausschließlich für eine Subgruppe (z. B. Erstgebärende, Gebärende mit Spontangeburt, spontaner Wehentätigkeit und PDA) oder eine größere Population (alle lebend geborenen Einlinge) Ergebnisse vorlagen und von unterschiedlichen Ergebnissen zwischen dieser Subgruppe und der Gesamtpopulation ausgegangen werden musste oder die untersuchte Expositionsgruppe nur eingeschränkt auf Deutschland übertragbar ist.

Schwerwiegende fehlende Genauigkeit führte zur Abwertung der Evidenz für einen Endpunkt, wenn die Ergebnisse zu unpräzise waren, um eine Entscheidung für oder gegen eine Empfehlung stützen zu können. Primär basierte eine Abwertung um 1 Kategorie auf Ergebnissen, deren KI sowohl einen relevanten Nutzen als auch Schaden einschließen. Eine Abwertung wegen fehlender Genauigkeit erfolgte um 2 Kategorien, wenn das KI des Effekts einen bedeutenden Nutzen und sehr großen Schaden einschließt. Es erfolgte keine Abwertung, wenn das KI des absoluten Effekts zwar sowohl einen Nutzen als auch Schaden einschließt, dieser Unterschied aufgrund der Effektgröße aber als nicht relevant eingestuft wurde.

Es konnte für keinen der untersuchten Endpunkte ein Hinweis auf einen Publikationsbias festgestellt werden, sodass **keine Abwertung wegen des Verdachts auf Publikationsbias** erfolgte.

In einzelnen Studien traten Dosis-Wirkungs-Beziehungen auf, die den Zusammenhang zwischen einer abnehmenden Geburtenzahl pro Jahr in den Geburtskliniken und einer steigenden oder abnehmenden Häufigkeit des Endpunktes beschreiben. Die Qualität der Evidenz wurde für keinen Endpunkt heraufgestuft, da diese Beziehung nicht konsistent auf alle vorliegenden Studien zutraf und bei allen Beobachtungsstudien schwerwiegende Studienlimitationen auftraten.

8.2.1 Geburtsmodus

In den 2 eingeschlossenen Beobachtungsstudien lagen ausschließlich Informationen zu den Ausprägungen vaginal-operative Geburt (1 Studie), vaginale Geburt (1 Studie) und ungeplanter Kaiserschnitt (1 Studie) vor, wobei für jede Ausprägung Ergebnisse aus maximal 2 Beobachtungsstudien vorlagen. Die in 1 Studie berichteten Ergebnisse zur vaginalen Geburt fassen die Ausprägungen Spontangeburt und vaginal-operative Geburt zusammen. Für den Kaiserschnitt werden die Ergebnisse für den ungeplanten Kaiserschnitt und nicht für einen geplanten Kaiserschnitt bewertet, da die Entscheidung für einen geplanten Kaiserschnitt per Definition in der Regel bereits vor Beginn der aktiven EP getroffen wird und somit Frauen mit einem geplanten Kaiserschnitt nicht der Zielpopulation dieses Evidenzberichts entsprechen.

Vaginal-operative Geburt

In 1 Studie wurden für eine gemeinsame Auswertung von Erst- und Mehrgebärenden ausschließlich relative Häufigkeiten für das Auftreten einer vaginal-operativen Geburt in Geburtskliniken mit größerer Geburtenzahl im Vergleich zu Geburtskliniken mit geringerer

Geburtenzahl berichtet: Insgesamt 16,5 % der Gebärenden in größeren und 12,7 % in Kliniken mit einer geringeren Geburtenzahl wiesen eine vaginal-operative Geburt auf. In dieser Studie beschreibt eine Dosis-Wirkungs-Beziehung eine sinkende Anzahl vaginal-operativer Geburten mit abnehmender Geburtenzahl in den Geburtskliniken.

Der primär interessierende Vergleich zwischen Geburtskliniken mit größeren und geringeren Geburtenzahlen konnte einen signifikanten Effekt nachweisen (≥ 3000 vs. < 1000 : OR: 1,47; 98,3 %-KI: [1,03; 2,09]).

Ausgehend von einem Risiko von 13 vaginal-operativen von 100 Geburten bei Erst- und Mehrgebärenden bei einer geringeren Erfahrung / Expertise, gemessen an der Geburtenzahl der Geburtsklinik, würden 5 (von 0 bis 11) mehr Gebärende eine vaginal-operative Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer größeren Geburtenzahl gebären.

Vaginale Geburt

In 1 Studie, in welche ausschließlich Erstgebärende eingeschlossen wurden, gebären in Kliniken mit einer großen, mittelgroßen und kleinen Geburtenzahl pro Jahr jeweils 88,6 %, 87,8 % und 87,5 % der 41 689, 27 970 und 4343 ausgewerteten Gebärenden vaginal. Es wurde ein signifikanter Unterschied zwischen Geburtskliniken mit größeren und mittelgroßen Geburtenzahlen (> 3000 vs. 1100–3000: OR: 1,22; 97,5 %-KI: [1,14; 1,30]) berichtet. Ausgehend von einem Risiko von 88 vaginalen pro 100 Geburten bei Erstgebärenden bei einer geringeren Erfahrung / Expertise in einer mittelgroßen Klinik, gemessen an der Geburtenzahl der Geburtsklinik, würden 2 (von 1 bis 3) mehr Gebärende eine vaginale Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer größeren Geburtenzahl gebären. Im Vergleich mittelgroßer und kleiner Geburtskliniken (1100–3000 vs. < 1100 Geburten pro Jahr) konnte kein Unterschied festgestellt werden.

Kaiserschnitt

Ungeplante Kaiserschnitte im Geburtsverlauf traten in 1 Studie in Geburtskliniken mit großen, mittelgroßen und kleinen Geburtenzahlen pro Jahr bei 11,4 %, 12,2 % und 12,5 % bei 41 689, 27 970 und 4343 Gebärenden auf.

In 1 der 2 primär interessierenden Vergleiche zwischen Geburtskliniken mit größeren und geringeren Geburtenzahlen werden für Erstgebärende signifikante Effekte berichtet (> 3000 vs. 1100–3000: OR: 0,82; 98,3 %-KI: [0,77; 0,88]).

Daher würden ausgehend von einem Risiko von 12 ungeplanten Kaiserschnitten von 100 Geburten bei Erstgebärenden bei einer geringeren Erfahrung / Expertise in einer mittelgroßen Klinik, gemessen an einer mittleren Geburtenzahl der Geburtsklinik, 2 (von 1 bis 3) weniger Gebärende einen ungeplanten Kaiserschnitt aufweisen, wenn sie in einer Geburtsklinik mit einer größeren Geburtenzahl gebären würden. Im Vergleich mittelgroßer und kleinerer Geburtskliniken konnte kein Unterschied festgestellt werden.

Zusammenfassung

Die Qualität der Evidenz aus den vorliegenden Beobachtungsstudien wurde für den Endpunkt Geburtsmodus zusammenfassend als sehr niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 18). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasste eine unklare Verblindung der Gebärenden und behandelnden Personen, eine unklare ergebnisunabhängige Berichterstattung und in 1 Studie eine möglicherweise unterschiedliche Erfassung von Risikofaktoren. Die Abstufung wegen schwerwiegender Indirektheit bei Mehrgebärenden liegt darin begründet, dass für die berichteten Ausprägungen des Geburtsmodus vaginale Geburt und Kaiserschnitt ausschließlich Informationen zu Erstgebärenden vorliegen. Darüber hinaus bleibt unklar, ob die berichteten Veränderungen des Geburtsmodus mit einem relevanten Nutzen oder einen Schaden auf weitere kritische Endpunkte verbunden sind.

Für die 3 berichteten Ausprägungen des Endpunkts Geburtsmodus konnten bei sehr niedriger Qualität der Evidenz auf Grundlage der sehr hohen berichteten Fallzahlen und der daraus folgenden hohen Präzision der Ergebnisse für mehrere berichtete Vergleiche statistisch signifikante Effekte festgestellt werden. Es konnte ein Einfluss der Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, nachgewiesen werden: In größeren Kliniken stieg die Anzahl vaginal-operativer Geburten, vaginaler Geburten bei Erstgebärenden (Spontangeburt und vaginal-operative Geburten) und die Anzahl der Kaiserschnitte bei Erstgebärenden sank.

8.2.2 Neonatale Mortalität

In den 2 eingeschlossenen Beobachtungsstudien wurden Informationen zur neonatalen Mortalität innerhalb von 27 Tagen nach der Geburt, zur frühen neonatalen Mortalität innerhalb von 7 Tagen nach der Geburt und zur intrapartalen Mortalität in jeweils 1 Studie berichtet.

In 1 Studie wurden keine absoluten und relativen Häufigkeiten, sondern ausschließlich Effektschätzer für das Auftreten neonataler Todesfälle bei 638 183 Gebärenden in großen Geburtskliniken und 31 745 Gebärenden in kleinen Geburtskliniken pro Jahr berichtet.

1 weitere Studie berichtet, dass in Universitätskliniken 0,24 % der Feten intrapartal und weitere 0,02 % der Neugeborenen von 92 139 Gebärenden innerhalb von 7 Tagen verstarben, während in kleineren Geburtskliniken 0,27 % der Feten intrapartal und weitere 0,03 % der Neugeborenen von 39 385 Gebärenden innerhalb von 7 Tagen verstarben. Eine Dosis-Wirkungs-Beziehung beschreibt eine steigende Anzahl intrapartaler Todesfälle mit abnehmender Geburtenzahl in den Geburtskliniken, wobei kein signifikanter Unterschied zwischen kleinen und großen Geburtskliniken nachgewiesen wurde.

Für keinen der 3 primär interessierenden Vergleiche zwischen Geburtskliniken mit größeren und geringeren Geburtenzahlen wurde ein signifikanter Effekt berichtet.

Die Qualität der Evidenz wurde für den Endpunkt neonatale Mortalität zusammenfassend als sehr niedrig bewertet, da aufgrund sehr schwerwiegender Studienlimitationen und sehr schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 18). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine fehlende Verblindung der Gebärenden und behandelnden Personen, eine unklare adäquate Confounderadjustierung bei fehlenden Informationen zu Baseline-Charakteristika in den Expositions- und Referenzgruppen, eine unklare ergebnisunabhängige Berichterstattung und der Ausschluss von Neugeborenen aufgrund fehlender Daten aus der Analyse, der dieses seltene Ereignis beeinflusst haben kann. Da zwar Ergebnisse aus 2 Studien vorliegen, diese aber unterschiedliche Zeitpunkte für das Versterben der Neugeborenen berichten, konnten die Ergebnisse der Studien nicht zusammengefasst werden und keine Bewertung der Inkonsistenz konnte vorgenommen werden. Die Abwertung wegen Indirektheit erfolgte, da das Ergebnis auf Informationen zu allen lebend geborenen Einlingen basiert und sich das Mortalitätsrisiko der vor der abgeschlossenen 37. SSW geborenen von den termingerecht geborenen Neugeborenen unterscheidet. Die Abwertung aufgrund schwerwiegender fehlender Genauigkeit erfolgte, da die berichteten Ergebnisse sowohl einen bedeutenden Nutzen als auch sehr großen Schaden umfassen.

Damit konnte für die neonatale Mortalität kein Einfluss der Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, nachgewiesen werden.

8.2.3 Schwere metabolische Acidose (pH < 7,0)

In 1 Beobachtungsstudie trat bei 0,39 % der 92 139 im Universitätskrankenhaus geborenen Neugeborenen und bei 0,25 % der 39 385 in kleineren Geburtskliniken mit ≤ 999 Geburten im Jahr geborenen Neugeborenen eine schwere metabolische Acidose auf. Es wurde ein signifikanter Unterschied mit einer höheren Anzahl von Neugeborenen mit einer schweren metabolischen Acidose im Universitätskrankenhaus berichtet (OR: 1,43; 95 %-KI: [1,14; 1,79]). Ausgehend von einem Risiko von 25 von 10 000 Neugeborenen mit schwerer metabolischer Acidose bei einer geringeren Erfahrung/Expertise, gemessen an der Geburtenzahl der Geburtsklinik, würden 11 (von 3 bis 20) mehr Neugeborene eine schwere metabolische Acidose aufweisen, wenn sie in einem Universitätskrankenhaus geboren würden.

Die Qualität der Evidenz wurde für den Endpunkt schwere metabolische Acidose zusammenfassend als sehr niedrig bewertet, da aufgrund sehr schwerwiegender Studienlimitationen herabgestuft wurde (siehe Tabelle 18). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine unklare Verblindung der Gebärenden und behandelnden Personen, fehlende Informationen zu Baseline-Charakteristika in den Expositions- und Referenzgruppen und eine unklare ergebnisunabhängige Berichterstattung. Zusätzlich wurden 324 (0,12 %) Neugeborene aufgrund fehlender Daten aus der Analyse ausgeschlossen, was das Ergebnis für dieses seltene Ereignis beeinflusst haben kann.

Damit konnte für das Auftreten einer schweren metabolischen Acidose ein Einfluss der Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, nachgewiesen werden. Alle Ergebnisse basieren auf den in den Datenbanken der Perinatalerhebung und Geburtsregister gespeicherten Informationen von Gebärenden mit geringem Risiko für Komplikationen. Die höhere Anzahl von Neugeborenen mit einer schweren metabolischen Acidose im Universitätskrankenhaus kann einerseits damit zusammenhängen, dass Schwangere mit einem höheren Geburtsrisiko in eine größere und gegebenenfalls spezialisierte Geburtsklinik überwiesen werden und so aufgrund risikoreicher Fälle die Häufigkeit von schweren metabolischen Acidosen steigt. Eine vollständige Berücksichtigung dieses Selektionsbias auf Grundlage der in den Datenbanken erfassten Informationen ist ohne Randomisierung nicht möglich. Andererseits könnte ein Zusammenhang zwischen größeren Geburtskliniken und einem interventionsreicheren Vorgehen im Vergleich zu kleineren Geburtskliniken, z. B. aufgrund unterschiedlicher Ausstattung oder personeller Besetzung, bestehen.

8.2.4 Episiotomie

Zum Endpunkt Episiotomie liegen verwertbare Informationen aus 1 Beobachtungsstudie vor.

In dieser Studie trat eine Episiotomie bei 17,0 % der Gebärenden in Geburtskliniken mit größeren Geburtenzahlen im Vergleich zu 19,9 % der Gebärenden in Geburtskliniken mit geringeren Geburtenzahlen auf. Damit konnte für den Einsatz von Episiotomien kein Einfluss der Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, nachgewiesen werden (OR: 0,80; 95 %-KI: [0,57; 1,12]).

Die Qualität der Evidenz wurde für den Endpunkt Episiotomie zusammenfassend als sehr niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen, schwerwiegender Indirektheit und schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 18). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine unklare Verblindung der Gebärenden und behandelnden Personen und eine unklare ergebnisunabhängige Berichterstattung. Die Abstufung wegen Indirektheit erfolgte, da in die Studie nur Frauen mit einer Spontangeburt und nicht Frauen mit einer vaginal-operativen Geburt eingeschlossen wurden, sodass die in die Studie eingeschlossene Population nur eine Teilpopulation der Gebärenden aus Deutschland abbildet. Die weitere Herabstufung basierte auf Ungenauigkeiten der Effektschätzung. Das 95 %-KI des relativen Risikos (OR: 0,80; 95 %-KI: [0,57; 1,12]) schließt sowohl einen relevanten Nutzen als auch einen Schaden ein.

8.2.5 Wehenmittel

Es liegen Informationen zum Endpunkt Wehenmittel aus 1 Beobachtungsstudie vor, in der Informationen zur Wehenmittelgabe während der Geburt berichtet werden. 62,0 % der Gebärenden in Geburtskliniken mit einer höheren Geburtenzahl erhielten im Vergleich zu 76,0 % der Gebärenden in Geburtskliniken mit einer geringeren Geburtenzahl während der

Geburt Oxytocin. Es wurde ein signifikanter Unterschied mit einer geringeren Anzahl an Wehenmittelgabe während der Geburt bei Gebärenden, die in größeren Geburtskliniken ≥ 4000 Geburten gebären, berichtet (OR: 0,67; 95 %-KI: [0,56; 0,83]). Ausgehend von einem Risiko für eine Wehenmittelgabe von 76 von 100 Gebärenden mit einer spontan eingesetzten Wehentätigkeit bei einer geringeren Erfahrung / Expertise, gemessen an einer geringeren Geburtenzahl der Geburtsklinik, würden 8 (von 4 bis 12) weniger Gebärende eine Wehenmittelgabe während der Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer größeren Geburtenzahl gebären würden. Diese Raten liegen deutlich über dem deutschen Populationsrisiko für Wehenmittelgaben unter der Geburt nach den Angaben des IQTIG [5]. Demnach erhielten 2017 in Deutschland bezogen auf alle Geburten 26 pro 100 Frauen Wehenmittel während der Geburt. Ausgehend von diesem Risiko für eine Wehenmittelgabe, würden 7 (von 3 bis 10) weniger Gebärende eine Wehenmittelgabe während der Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer sehr großen Geburtenzahl gebären würden.

Die Qualität der Evidenz wurde für den Endpunkt Wehenmittel zusammenfassend als sehr niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 18). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine unklare Verblindung der Gebärenden und behandelnden Personen sowie eine unklare ergebnisunabhängige Berichterstattung. In die Studie wurden nur Frauen mit einer für Frankreich typischen Standardbetreuung eingeschlossen, die eine spontane Wehentätigkeit und eine PDA-Anlage umfasste, und es liegen ausschließlich Ergebnisse für einen Vergleich zu sehr großen Kliniken mit > 4000 Geburten pro Jahr vor, welche in Deutschland sehr selten sind. In Deutschland erhielten rund 23 % aller Gebärenden während der Geburt eine PDA [5], sodass die in die Studie eingeschlossene Population nur eine Teilpopulation der Gebärenden aus Deutschland abbildet, weswegen die Abstufung wegen Indirektheit erfolgte.

8.2.6 Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte

Es konnten Effektschätzungen für insgesamt 3 kritische Endpunkte (Geburtsmodus, neonatale Mortalität, schwere metabolische Acidose [(pH $< 7,0$)] und 2 wichtige Endpunkte (Episiotomie, Wehenmittel) auf der Grundlage von Beobachtungsstudien ermittelt werden. Die Qualität der Evidenz wurde für alle berichteten Endpunkte als sehr niedrig bewertet.

Es wurden keine Ergebnisse zur maternalen Mortalität, höhergradigen Dammrissen (III^o, IV^o), hohem mütterlichen Blutverlust ≥ 1000 ml, zur Aufnahme einer Stillbeziehung bei initialem Stillwunsch, zur teilweisen / vollständigen Plazentaretention, zum Geburtserleben, zur subjektiven Gesundheitseinschätzung, zum Apgar-Score nach 5 Minuten < 6 , zum Basendefizit ≥ -16 mmol/l und zur Verlegung auf die Neugeborenen-Intensivstation und zu weiteren wichtigen Endpunkten berichtet.

Zusammenfassend wurde die Qualität der Evidenz als sehr niedrig bewertet, sodass folglich das Gesamtvertrauen in die Effektschätzungen als sehr niedrig bewertet wird und der wahre Effekt

einer größeren Erfahrung / Expertise, beschrieben durch die Geburtenzahlen pro Jahr in den Geburtskliniken, wahrscheinlich relevant verschieden von der berechneten Effektschätzung ist.

8.3 Beschreibung des Evidenzprofils zur Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe)

Im Folgenden wird die vorliegende Evidenz zu dem 1 berichteten kritischen Endpunkt der eingeschlossenen Studien zur Versorgungsstufe der Geburtsklinik zusammenfassend dargestellt.

Die Angabe zum absoluten Effekt basiert auf dem Risiko der Referenzgruppe mit niedriger Versorgungsstufe, da Angaben der Bundesauswertung zur stationären Qualitätssicherung zum Erfassungsjahr 2017 des Instituts für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) [5] keine Differenzierungsmöglichkeit zwischen den verschiedenen Versorgungsstufen der Geburtskliniken enthalten. Ausgehend von einer niedrigen Qualität der Evidenz, bedingt durch das Studiendesign der eingeschlossenen Studie, erfolgte eine Herabstufung aufgrund sehr schwerwiegender Studienlimitationen und schwerwiegender fehlender Genauigkeit.

Schwerwiegende Studienlimitationen mit der Möglichkeit einer Herabstufung der Qualität der Evidenz um 2 Kategorien umfassten

- eine fehlende Verblindung der Teilnehmerinnen und behandelnden Personen,
- eine unklare ergebnisunabhängige Berichterstattung, da Angaben zu geplanten primären und sekundären Endpunkten oder eine Differenzierung diesbezüglich fehlen und kein Studienprotokoll vorliegt.

Schwerwiegende fehlende Genauigkeit ermöglicht die Herabstufung der Qualität der Evidenz für einen Endpunkt, wenn die Ergebnisse zu unpräzise waren, um eine Entscheidung für oder gegen eine Empfehlung stützen zu können. Primär basierte eine Abwertung um 1 Kategorie auf einem KI, das sowohl einen relevanten Nutzen als auch Schaden einschließen kann.

Es erfolgte **eine Abwertung wegen Indirektheit**, wenn ausschließlich für eine Subgruppe (z. B. Gebärende mit spontaner Wehentätigkeit und PDA) Ergebnisse vorlagen und von unterschiedlichen Ergebnissen zwischen dieser Subgruppe und der Gesamtpopulation ausgegangen werden musste.

Da die Ergebnisse auf nur 1 Studie basieren, erfolgte **keine Abwertung wegen Inkonsistenz**, und die Inkonsistenz wurde als nicht relevant beschrieben.

Es konnte kein Hinweis auf einen Publikationsbias festgestellt werden, sodass **keine Abwertung wegen des Verdachts auf Publikationsbias** erfolgte.

8.3.1 Geburtsmodus

Angaben zu absoluten Zahlen liegen in der 1 vorliegenden Beobachtungsstudie nur für Spontangeburt vor. Insgesamt traten in der Studie 7122 Spontangeburt bei insgesamt 9530 eingeschlossenen Gebärenden auf, wobei nicht für alle Ausprägungen des Endpunktes Geburtsmodus Informationen zu allen eingeschlossenen Frauen vorliegen. Die vorliegenden Informationen zum Kaiserschnitt wurden nicht extrahiert, da die Gesamtanzahl der Kaiserschnitte ohne Differenzierungsmöglichkeit zwischen einem geplanten und ungeplanten Kaiserschnitt berichtet wird. Zu den vaginal-operativen Geburten liegen ausschließlich relative Häufigkeitsangaben vor. 15 % der Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer größeren Versorgungsstufe (Level 3) und 15,9 % der Gebärenden in der Gruppe mit einer Betreuung in einer Geburtsklinik mit einer geringeren Versorgungsstufe (Level 1) wiesen eine vaginal-operative Geburt auf.

Die Qualität der Evidenz wurde für den Endpunkt Geburtsmodus zusammenfassend als sehr niedrig bewertet, da aufgrund sehr schwerwiegender Studienlimitationen und schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 19). Die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfasst eine unklare Verblindung der teilnehmenden und behandelnden Personen, fehlende Informationen zu Baseline-Charakteristika in der Expositions- und Referenzgruppe, eine fehlende Verblindung von Gebärenden und behandelnden Personen und eine unklare ergebnisunabhängige Berichterstattung.

Für die 1 Ausprägung des Endpunktes Geburtsmodus konnte für das Merkmal Versorgungsstufe kein signifikanter Einfluss auf die untersuchten Vergleichsgruppen für vaginal-operative Geburten festgestellt werden.

Die Berechnung des absoluten Effekts basiert auf dem Risiko der Referenzgruppe mit geringer Versorgungsstufe, in welcher 16 von 100 Geburten vaginal-operativ erfolgten.

Für den Geburtsmodus konnte für die Anzahl von vaginal-operativen Geburten durch eine größere Erfahrung / Expertise, beschrieben durch die Versorgungsstufe der Geburtsklinik, kein Unterschied nachgewiesen werden. Es fehlen Informationen zu den beiden Ausprägungen Spontangeburt und Kaiserschnitt, um den Einfluss einer größeren Erfahrung / Expertise auf den Endpunkt Geburtsmodus bewerten zu können.

8.3.2 Episiotomie

Es liegen Informationen zum Endpunkt Episiotomie aus 2 Beobachtungsstudien vor. Da die eingeschlossenen Frauen aus 1 Studie sich nur auf Frauen mit einer Spontangeburt beziehen, die eine Subgruppe der vaginal gebärenden Frauen der 2. Studie bilden, wird im Evidenzprofil das Ergebnis für die Frauen mit einer vaginalen Geburt dargestellt, da diese Daten die Gesamtpopulation besser abbilden.

In der Studie trat eine Episiotomie bei 24,4 % der Gebärenden in Geburtskliniken mit einer höheren Versorgungsstufe im Vergleich zu 27,6 % der Gebärenden in Geburtskliniken mit einer geringeren Versorgungsstufe auf. Damit konnte für den Einsatz von Episiotomien kein Einfluss der Erfahrung / Expertise, gemessen an einer höheren Versorgungsstufe, nachgewiesen werden (RR: 1,11; 95 %-KI: [0,91; 1,36]).

Die Qualität der Evidenz wurde für den Endpunkt Episiotomie zusammenfassend als sehr niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender fehlender Genauigkeit herabgestuft wurde (siehe Tabelle 19). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine unklare Verblindung der Gebärenden und behandelnden Personen und eine unklare ergebnisunabhängige Berichterstattung. Die weitere Herabstufung basierte auf Ungenauigkeiten der Effektschätzung. Das 95 %-KI des relativen Risikos (RR: 1,11; 95 %-KI: [0,91; 1,36]) schließt sowohl einen relevanten Nutzen als auch einen Schaden ein.

8.3.3 Wehenmittel

Es liegen Informationen zum Endpunkt Wehenmittel aus 1 Beobachtungsstudie vor, in der Informationen zur Wehenmittelgabe während der Geburt berichtet werden. 66,0 % der Gebärenden in Geburtskliniken mit einer höheren Versorgungsstufe erhielten im Vergleich zu 75,0 % der Gebärenden in Geburtskliniken mit einer geringeren Versorgungsstufe während der Geburt Oxytocin. Es wurde ein signifikanter Unterschied mit einer geringeren Anzahl an Wehenmittelgabe während der Geburt bei Gebärenden, die in größeren Geburtskliniken mit einer höheren Versorgungsstufe gebären, berichtet (OR: 0,63; 95 %-KI: [0,51; 0,77]). Ausgehend von einem Risiko für eine Wehenmittelgabe während der Geburt von 75 von 100 Gebärenden mit einer spontan eingesetzten Wehentätigkeit bei einer geringeren Erfahrung / Expertise, gemessen an einer geringeren Versorgungsstufe der Geburtsklinik, würden 10 (von 5 bis 15) weniger Gebärende eine Wehenmittelgabe während der Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer höheren Versorgungsstufe gebären würden. Diese Raten liegen deutlich über dem deutschen Populationsrisiko für Wehenmittellgaben unter der Geburt nach den Angaben des IQTIG [5]. Demnach erhielten 2017 in Deutschland bezogen auf alle Geburten 26 pro 100 Frauen Wehenmittel während der Geburt. Ausgehend von diesem Risiko für eine Wehenmittelgabe, würden 8 (von 5 bis 11) weniger Gebärende eine Wehenmittelgabe während der Geburt aufweisen, wenn sie in einer Geburtsklinik mit einer höheren Versorgungsstufe gebären würden.

Die Qualität der Evidenz wurde für den Endpunkt Wehenmittel zusammenfassend als sehr niedrig bewertet, da aufgrund schwerwiegender Studienlimitationen und schwerwiegender Indirektheit herabgestuft wurde (siehe Tabelle 19). Die Gründe für die Abstufung wegen Studienlimitationen, die das Vertrauen in die Effektschätzungen beeinträchtigen, umfassten eine unklare Verblindung der Gebärenden und behandelnden Personen und eine unklare ergebnisunabhängige Berichterstattung. In die Studie wurden nur Frauen mit einer für Frankreich typischen Standardbetreuung eingeschlossen, die eine spontane Wehentätigkeit und

eine PDA-Anlage umfasste. In Deutschland erhalten rund 23 % aller Gebärenden während der Geburt eine PDA [5], sodass die in die Studie eingeschlossene Population nur eine Teilpopulation der Gebärenden aus Deutschland abbildet, weswegen die Abstufung wegen Indirektheit erfolgte.

8.3.4 Gesamtbeurteilung der Qualität der Evidenz für alle Endpunkte

Es konnten Effektschätzungen für insgesamt 1 Ausprägung (vaginal-operative Geburt) von 1 kritischen Endpunkt (Geburtsmodus) und 2 wichtigen Endpunkten (Episiotomie, Wehenmittel) auf der Grundlage von 2 Beobachtungsstudien ermittelt werden, für die die Qualität der Evidenz als sehr niedrig bewertet wurde. Es wurden keine Ergebnisse zu maternaler und neonataler Mortalität, höhergradigen Dammrissen (III°, IV°), hohem mütterlichen Blutverlust ≥ 1000 ml, zur Aufnahme einer Stillbeziehung bei initialem Stillwunsch, zur teilweisen / vollständigen Plazentaretention, zum Geburtserleben, zur subjektiven Gesundheitseinschätzung, zum Apgar-Score nach 5 Minuten < 6 , zu schwerer metabolischer Acidose ($\text{pH} < 7,0$), zum Basendefizit ≥ -16 mmol/l und zur Verlegung auf die Neugeborenen-Intensivstation und zu weiteren wichtigen Endpunkten berichtet.

Zusammenfassend wurde die Qualität der Evidenz als sehr niedrig bewertet, sodass folglich das Gesamtvertrauen in die Effektschätzung als sehr niedrig bewertet wird und der wahre Effekt einer größeren Erfahrung / Expertise beschrieben durch die Versorgungsstufe der Geburtsklinik wahrscheinlich relevant verschieden von der berechneten Effektschätzung ist.

9 Diskussion

Ziel dieses Rapid Reports war die zusammenfassende Darstellung und Bewertung der vorliegenden Evidenz zum Einfluss von Erfahrung und Expertise der ärztlichen Geburtshelferinnen und Geburtshelfer, Hebammen und Entbindungspfleger auf die vorgegebenen maternalen und neonatalen Endpunkte. Die Erfahrung / Expertise einer einzelnen Person oder einer Geburtsklinik ist nicht anhand einer objektiven Größe messbar. Der Bewertung kann sich nur über messbare Faktoren genähert werden, die als Indikator zur Messung von Erfahrung / Expertise verwendet werden. Faktoren, welche verschiedene Ausprägungen von Erfahrung und Expertise beschreiben können, wurden im Vorfeld von der Leitliniengruppe festgelegt und umfassten individuelle Faktoren des geburtshilflichen Personals (Anstellungsverhältnis, Berufsjahre / Weiterbildung, Anzahl von Geburten pro Fachperson, Qualifikation) und systemimmanente Faktoren (Krankenhäuser mit unterschiedlichen Geburtenzahlen, Geburtskliniken mit angeschlossener Kinderklinik oder kooperierenden Kinderkliniken).

Da das Wort *Erfahrung* „bestimmte Kenntnisse oder Einsichten, zu denen jemand durch meist wiederholte Wahrnehmung gelangt ist“ [62] beinhaltet, ist die theoretische Annahme, dass eine Fachperson mit einer größeren Anzahl an von ihr begleiteter Geburten über eine größere Erfahrung / Expertise verfügt, plausibel. Eine große Anzahl begleiteter Geburten kann z. B. durch eine große Anzahl an Berufsjahren oder eine Vollzeit-Anstellung erreicht werden. Diese beiden Faktoren können auf eine hohe Anzahl an Geburten pro Fachperson hinweisen, garantieren dies jedoch nicht. So kann eine in Vollzeit angestellte Fachperson, die in einer Geburtsklinik mit einer vergleichsweise geringen Geburtenzahl pro Jahr arbeitet, gemessen an der Anzahl von Geburten pro Fachperson, eine geringere Erfahrung / Expertise aufweisen als eine in Teilzeit angestellte Fachperson, die in einer Geburtsklinik mit einer sehr hohen Geburtenzahl arbeitet. Ebenso kann Erfahrung / Expertise in Form von Fachkenntnissen oder speziellem Wissen, z. B. durch Weiterbildungsmaßnahmen, erlangt werden. Individuelle fachliche Expertise kann somit sowohl über Berufserfahrung, die „die in langen Jahren gesammelte Erfahrung im Beruf“ [62] beschreibt, als auch durch Qualifikations- und Weiterbildungsmaßnahmen erworben werden, auch wenn sich das so erworbene Wissen, die sogenannte Sekundärerfahrung, durch die fehlende wiederholte Wahrnehmung von der eigenen Erfahrung unterscheidet.

Die systemimmanenten Faktoren beziehen sich auf die Größe der Krankenhäuser gemessen an den Geburtenzahlen pro Jahr und dem Versorgungsgrad der Geburtskliniken. In Deutschland werden Geburtskliniken entsprechend ihren verbindlichen Mindestanforderungen an die Versorgung von Schwangeren und Neugeborenen in 4 Stufen eingeteilt (Versorgungsstufe I: Perinatalzentrum Level 1, Versorgungsstufe II: Perinatalzentrum Level 2, Versorgungsstufe III: Perinataler Schwerpunkt, Versorgungsstufe IV: Geburtsklinik) [63]. Größere Geburtenzahlen pro Jahr einer Geburtsklinik müssen nicht zwangsläufig für eine höhere Versorgungsstufe sprechen, sind jedoch in der Regel ein guter Indikator. Die Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen zur Qualitätssicherung der Versorgung von Früh- und Reifgeborenen [63] beschreibt Anforderungen an die Besetzung des

geburtshilflichen Fachpersonals in den Geburtskliniken. So müssen Geburtskliniken der Versorgungsstufen I und II eine 24-Stunden-Präsenz einer Ärztin / eines Arztes vor Ort gewährleisten und eine Fachärztin / ein Facharzt muss mindestens im Hintergrund erreichbar sein. Ebenso muss die 24-Stunden-Präsenz von 1 Hebamme oder 1 Entbindungspfleger gegeben sein und 1 weitere Hebamme muss mindestens in Rufbereitschaft zur Verfügung stehen. Für die Versorgungsstufen III und IV befinden sich in der Richtlinie keine Vorgaben zur personellen Besetzung des geburtshilflichen Fachpersonals, weswegen in Geburtskliniken der geringeren Versorgungsstufen unter anderem keine 24-Stunden-Präsenz des geburtshilflichen Fachpersonals zwingend erforderlich ist. Kritisch an der Beschreibung der Erfahrung / Expertise anhand der Faktoren Geburtenzahl oder der Versorgungsstufe der Geburtsklinik ist, dass die entsprechend der Definition von Erfahrung vorgegebene Verknüpfung mit einer bestimmten Person fehlt. So kann eine Berufsanfängerin / ein Berufsanfänger mit einer geringen individuellen Erfahrung / Expertise in einer Geburtsklinik mit einer an systemimmanenten Faktoren gemessenen großen Erfahrung / Expertise arbeiten und umgekehrt eine Person mit einer großen Berufserfahrung und somit individuellen hohen Erfahrung / Expertise in einer Geburtsklinik mit einer vergleichsweise geringen Erfahrung / Expertise gemessen an der Geburtenzahl oder der Versorgungsstufe der Geburtsklinik. Eine ausschließliche Betrachtung systemimmanenter Faktoren berücksichtigt nicht den Einfluss der individuellen Erfahrung / Expertise.

2 eingeschlossene RCTs (Lomas 1991, Chaillet 2015) untersuchen die Auswirkung einer durch Qualifikations- und Weiterbildungsmaßnahmen verbesserten individuellen fachlichen Expertise auf die vorgegebenen Endpunkte. Ziel beider Studien war die Reduktion der auch in Deutschland seit Jahren steigenden Kaiserschnittraten [64]. Dabei untersuchte 1 der eingeschlossenen Studien (Lomas 1991) bereits in den Jahren 1988 bis 1989 in kanadischen Krankenhäusern 3 Weiterbildungsmaßnahmen zur Umsetzung einer Leitlinie, welche die Betreuung von Frauen mit einem vorherigen Kaiserschnitt thematisierte. Ein vorheriger Kaiserschnitt oder eine andere Uterus-OP wurden 2017 in Deutschland bei 19,8 % der Kaiserschnitte als alleinige Indikation angegeben [5], sodass die Erhöhung der Erfahrung / Expertise in Form von Weiterbildungen für die Betreuung von Gebärenden mit einem vorherigen Kaiserschnitt von großer Bedeutung ist. In der eingeschlossenen Studie führte die Schulung von Meinungsführern zu einer signifikanten Erhöhung der vaginalen Geburten bei einer abnehmenden Anzahl der geplanten und ungeplanten Kaiserschnitte, ohne dass Auswirkungen auf die berichteten neonatalen Endpunkte (Todgeburten, Neugeborene mit Apgar-Score < 7) nachgewiesen werden konnten. Die Ergebnisse dieser Studie sind allerdings nur sehr eingeschränkt auf das deutsche Versorgungssystem übertragbar. Seit Ende der 1980er-Jahre, in welcher die Studie durchgeführt wurde, hat sich die Versorgung Gebärender verändert, was sich zum Beispiel im Anstieg der Kaiserschnittrate zeigt [64]. Diese Veränderungen und Indikationen für die Entscheidung für einen ungeplanten Kaiserschnitt wie Risikofaktoren für Komplikationen rund um die Schwangerschaft und Geburt, eine veränderte Risikobewertung in der Schwangerenbetreuung und Geburtshilfe und Veränderungen in der Klinikorganisation werden im Hintergrund dieses Berichtes beschrieben. Zum anderen ist das

Betreuungsverhältnis in US-amerikanischen Geburtskliniken aufgrund von Unterschieden in der Versorgungsstruktur, Qualifikation und Aufgabenverteilung des ärztlichen und pflegerischen Personals nur eingeschränkt auf das deutsche Versorgungsverhältnis übertragbar.

Ziel der zweiten, ebenfalls in Kanada in den Jahren 2008 bis 2011 durchgeführten Studie, in welcher Daten zu insgesamt 184 952 Gebärenden erhoben wurden, war die Untersuchung der Auswirkungen von auf Schulungsmaßnahmen und Audit- und Feedbackstrategien basierenden Weiterbildungsmaßnahmen zur Reduktion der Kaiserschnittraten bei Erst- und Mehrgebärenden. In Ergänzung zu Lomas 1991 wurden zahlreiche von der Leitliniengruppe als „kritisch für die Entscheidung“ oder „klinisch wichtig, aber für die Empfehlung nicht entscheidend“ kategorisierte Endpunkte untersucht. Die Ergebnisse dieser Studie können auf das deutsche Versorgungsverhältnis übertragen werden, sodass die Qualität der Evidenz zusammenfassend mit moderat bewertet wurde.

Zusammenfassend konnte nach Durchführung der Weiterbildungsmaßnahmen für Gebärende mit vaginal geplanter Geburt eine Zunahme der vaginalen Geburten (1 mehr [von 0 bis 3 mehr]), aber keine signifikante Abnahme der ungeplanten Kaiserschnittentbindungen (2 weniger [von 0 mehr bis 5 weniger]) nachgewiesen werden. Zu dem in den Studien gewählten Konfidenzniveau von 95 % konnte ausschließlich für eine Prüfintervention (Schulung von Meinungsführern in Lomas 1991), welche insgesamt 345 der 45 486 der in die Auswertung eingehenden Gebärenden betraf, eine Reduktion der ungeplanten Kaiserschnittentbindungen unter der Geburt nachgewiesen werden.

Da die Entscheidung für einen geplanten Kaiserschnitt bereits vor Geburtsbeginn erfolgt, erfolgte, soweit möglich, in diesem Evidenzbericht eine Darstellung der Ergebnisse für den ungeplanten Kaiserschnitt. Chaillet 2015 beobachtete nach Durchführung der Weiterbildungsmaßnahmen eine Abnahme der Gesamtanzahl aller (geplanten und ungeplanten) Kaiserschnitte um 1,8 Prozentpunkte (RD: -1,8; 95 %-KI: [-3,8; -0,2]). Bei einer Differenzierung zwischen geplanten und ungeplanten Kaiserschnitten konnte in Chaillet 2015 weder für geplante noch für ungeplante Kaiserschnitte ein signifikanter Unterschied zwischen den Interventionsgruppen nachgewiesen werden (RD: -1,0; 95 %-KI: [-3,2; 2,0] und RD: -0,9; 95 %-KI: [-2,1; 0,3]). In Lomas 1991 berichten die Autorinnen und Autoren einen signifikanten Unterschied im Auftreten geplanter Kaiserschnitte ($p = 0,001$). Diese traten bei 53,7 %, 69,7 % und 66,8 % der Gebärenden in den Interventionsgruppen mit Schulungen von Meinungsführern, Audit und Feedback und der Gruppe mit der Vergleichsintervention auf. Zusätzlich konnte in diesem Evidenzbericht für die neonatale Mortalität ein Vorteil der Weiterbildungsmaßnahmen nachgewiesen werden, auch wenn dieser auf den Neugeborenen aller Gebärenden, einschließlich derer mit geplantem Kaiserschnitt, basiert. Dieses Ergebnis entspricht den Ergebnissen von Chaillet 2015, welche für die dort berichteten zusammenfassenden Endpunkte eine Abnahme der schwerwiegenden und weniger schwerwiegenden neonatalen Morbidität nachweisen konnte (RD: -0,73; 95 %-KI: [-1,31; -0,08] und RD: -1,7; 95 %-KI: [-2,6; -0,9]). Der berichtete kombinierte Endpunkt „schwerwiegende neonatale Morbidität“ fasst die Komponenten intrapartale oder neonatale Todesfälle, Apgar-Score nach 5 Minuten < 4 ,

schwerwiegende metabolische Acidose ($\text{pH} < 7$), schwerwiegende Traumata, intraventrikuläre Hämorrhagie, Krämpfe innerhalb von 24 Stunden nach der Geburt, neurologische Schäden, invasive mechanische Beatmung, nekrotisierende Enterocolitis und hypoxisch-ischämische Enzephalopathie zusammen. Die Ergebnisse aller zusammengefassten Ereignisse werden einzeln berichtet. Damit fließt die in diesem Evidenzbericht beobachtete Zunahme der schweren metabolischen Acidose (RD: 0,04; 95 %-KI: [-0,06; 0,19]) in den Endpunkt schwerwiegende neonatale Morbidität ein. Der berichtete kombinierte Endpunkt „weniger schwerwiegende neonatale Morbidität“ fasst die Komponenten kardiopulmonare Morbidität, Apgar-Score nach 5 Minuten ≥ 4 und < 7 , fortgeschrittene metabolische Acidose ($\text{pH} \geq 7$ und $< 7,1$), weniger schwerwiegendes Trauma, nicht invasive mechanische Beatmung, Bluttransfusion und neonatale Infektion / Sepsis zusammen. Damit fließt die in diesem Evidenzbericht beobachtete Zunahme der fortgeschrittenen metabolischen Acidose (RD: 0,7; 95 %-KI: [0,0; 1,6]) in den Endpunkt weniger schwerwiegende neonatale Morbidität ein. Im Gegensatz zur neonatalen Morbidität konnte Chaillet 2015 für das Auftreten einer schwerwiegenden oder weniger schwerwiegenden maternalen Morbidität, zu denen die in diesem Evidenzbericht untersuchten Endpunkte höhergradige Dammriss, Infektion / Sepsis und die Mortalität gehören, keine Auswirkung der Weiterbildungsmaßnahmen feststellen. In diesem Evidenzbericht wird die Häufigkeit der Gabe von Wehenmitteln als wichtiger Endpunkt untersucht. Nach erfolgten Weiterbildungsmaßnahmen wurden Wehenmittel seltener zur Einleitung der Geburt (RD: -3,8; 95 %-KI: [-5,1; -2,7]), aber häufiger unter der Geburt RD: 3,2; 95 %-KI: [1,9; 4,6] verabreicht.

Beide eingeschlossenen Studien wurden auch in die aktuellste systematische Übersicht [65] zur Wirksamkeit nicht klinischer Interventionen zur Reduktion nicht notwendiger Kaiserschnittverbindungen eingeschlossen. Diese systematische Übersicht schloss sowohl RCTs als auch nicht randomisierte Studien, kontrollierte Studien mit Vorher-nachher-Vergleichen und Studien mit dem Interrupted Time Series Design ein. Die Interventionen richteten sich an Gebärende, das die Gebärenden im Zeitraum um die Geburt betreuende medizinische Personal und größere Organisationseinheiten. 2 der in die systematische Übersicht eingeschlossenen clusterrandomisierten Studien (Lomas 1991, Chaillet 2015) untersuchten Maßnahmen zur Verbesserung der Erfahrung / Expertise des medizinischen Fachpersonals und waren damit für diesen Evidenzbericht relevant.

Die weiteren eingeschlossenen Studien untersuchen ausschließlich systemimmanente Faktoren, die den Einfluss der Geburtenzahlen oder der Versorgungsstufe der Geburtskliniken auf maternale und neonatale Endpunkte untersuchen. Dabei sind die in den Studien gewählten Kategorien der Geburtenzahlen der Geburtskliniken sehr heterogen. In einigen Studien werden Kategorien in 1000er Schritten gebildet (z. B. Vergleich von Geburtskliniken mit > 3000 , < 3000 – 2000 , < 2000 – 1000 , < 1000 Geburten pro Jahr), während in anderen Studien zum Teil kleinere Kategorien gewählt werden (z. B. Vergleich von Geburtskliniken mit ≥ 1000 – 2499 , 500 – 999 und < 500 Geburten pro Jahr). In 1 Studie werden 3 Kategorien gebildet und alle Vergleiche erfolgten zur mittleren Kategorie mit 1100 – 3000 Geburten pro Jahr. Auch aufgrund

der teilweise sehr unterschiedlich gewählten Kategorien und Vergleiche fällt ein zusammenfassender Vergleich der Ergebnisse schwer, sodass in der Ergebnisdarstellung die primär interessierenden Vergleiche (größere im Vergleich zu geringerer Erfahrung / Expertise gemessen an der größten und kleinsten auf deutsche Verhältnisse übertragbaren Kategorie) herangezogen wurden.

Für die primär interessierenden Vergleiche konnte für die Anzahl vaginal-operativer Geburten, vaginaler Geburten und ungeplanter Kaiserschnitte, das Auftreten einer schweren metabolischen Acidose und Wehenmittel ein Einfluss der Erfahrung / Expertise, gemessen an größeren Geburtenzahlen pro Jahr in den Geburtskliniken, nachgewiesen werden. In größeren Geburtskliniken gebären mehr Erstgebärende vaginal, es traten weniger ungeplante Kaiserschnitte bei Erstgebärenden auf und Wehenmittel wurden bei Gebärenden mit einer PDA seltener verabreicht. In diesen Kliniken wurden aber auch mehr Kinder vaginal-operativ geboren und es wurde häufiger eine schwere metabolische Acidose diagnostiziert, die aber nicht mit einer höheren intrapartalen oder neonatalen Mortalität assoziiert war.

Die festgestellten höheren Komplikationsraten in Geburtskliniken mit größeren Geburtenzahlen basieren auf Ergebnissen aus 2 Studien: In 1 Studie (Coulm 2012) wurde eine höhere Anzahl von vaginal-operativen Geburten und in 1 weiteren Studie (Karalis 2016) ein häufigeres Auftreten einer schweren metabolischen Acidose festgestellt. Diese Ergebnisse könnten einerseits damit erklärt werden, dass Schwangere mit einem höheren Geburtsrisiko in der Regel in eine größere und gegebenenfalls spezialisierte Geburtsklinik überwiesen werden und dadurch die Anzahl invasiver oder interventionsreicher Geburten in diesen Kliniken steigt [6,56,59,66]. Veränderungen im Geburtsmodus mit einer höheren Anzahl vaginal-operativer Geburten können dadurch erklärt werden, dass bei auftretenden Komplikationen im Geburtsverlauf das geburtshilfliche Personal zwischen den Alternativen Kaiserschnitt oder vaginal-operativ Geburt entscheiden muss. Eine genauere Untersuchung der Auswirkung auf alle Ausprägungen des Geburtsmodus war nicht möglich, da interpretierbare Ergebnisse zu Spontangeburt und ungeplanten Kaiserschnitten im Geburtsverlauf nicht berichtet wurden. Laut Einschlusskriterien der eingeschlossenen Beobachtungsstudien sollten in beide Studien Schwangere mit einem geringen geburtshilflichen Risiko in die Auswertungen eingeschlossen werden. Die Autorinnen und Autoren in Karalis 2016 [57] berichten jedoch, dass in der Auswahl von Frauen mit einem geringen Risiko für Komplikationen nicht alle potenziellen maternalen und fetalen Risikofaktoren berücksichtigt werden konnten, da diese im Geburtsregister nicht vollständig erfasst wurden. Diese Feststellung wird auch von den Autorinnen und Autoren der am umfangreichsten adjustiertesten eingeschlossenen Studie dieses Evidenzberichtes als Limitation angemerkt [59]. Eine vollständige Berücksichtigung aller Risikofaktoren war aufgrund der fehlenden Randomisierung nicht möglich und unterschiedliche Risikoprofile der Gebärenden in den verschiedenen Expositionsgruppen können in nicht randomisierten Beobachtungsstudien nicht ausgeschlossen werden. Aus diesem Grund wird in der Bewertung der Qualität der Evidenz bei Beobachtungsstudien von einer niedrigen Qualität der Evidenz ausgegangen [22]. Dieser Selektionsbias konnte in der Studie,

auf der die Ergebnisse zur schweren metabolischen Acidose basieren, durch Confounder-adjustierung nicht vollständig ausgeglichen werden, sodass eine Verzerrung der Ergebnisse nicht ausgeschlossen werden kann. Wichtig ist, dass das häufigere Auftreten einer schweren metabolischen Acidose in 1 Studie (Karalis 2016) auf sehr geringen Unterschieden im Auftreten dieses seltenen Ereignisses basiert. Es war mit einer häufigeren Verlegung von Neugeborenen auf die Neugeborenen-Intensivstation, die Beobachtungsstation oder in ein anderes Krankenhaus assoziiert, aber nicht mit einer signifikanten Erhöhung der intrapartalen oder frühen neonatalen Mortalität innerhalb von 7 Tagen nach der Geburt.

Eine weitere Erklärung für die höhere Anzahl invasiver oder interventionsreicher Geburten in größeren Geburtskliniken wird von mehreren Autorinnen und Autoren darin gesehen, dass auch Schwangere mit einem geringen Risiko für Komplikationen in Geburtskliniken mit einer größeren Geburtenzahl eine invasivere Geburtshilfe erfahren als in Geburtskliniken mit einer geringeren Geburtenzahl [55,66]. Diese invasivere Geburtshilfe könnte auf unterschiedliche Ausstattungen oder die personelle Besetzung zurückgeführt werden. Diesen Zusammenhang zwischen Behandlungseffekten und Ergebnissen bei Schwangeren mit einem geringen Risikoprofil untersuchten Reif et al. 2017 [67] in einer prospektiv durchgeführten Kohortenstudie in österreichischen Geburtskliniken. Sie stellten fest, dass Neugeborene von Gebärenden mit einem in der Schwangerschaft festgestellten sehr geringen Risiko für Komplikationen sowohl in spezialisierten Perinatalzentren als auch bei einer Betreuung der Geburt durch eine erfahrene Ärztin / einen erfahrenen Arzt deutlich häufiger schwerwiegende Ereignisse aufwiesen als Gebärende in kleineren Kliniken oder bei einer Betreuung durch weniger spezialisiertes geburtshilfliches Personal. Dagegen profitierten Gebärende mit steigendem Risiko von einer Betreuung in einem spezialisierten Perinatalzentrum und einer Begleitung der Geburt durch eine erfahrene Ärztin / einen erfahrenen Arzt.

In größeren und spezialisierten Geburtskliniken wurden bei Frauen mit spontan einsetzender Wehentätigkeit und PDA deutlich seltener Wehenmittel im Vergleich zu kleineren Geburtskliniken verabreicht. Dieses Ergebnis basiert auf der französischen Perinataldatenerhebung im Jahr 2010. Dort erhielten 71 % der Gebärenden mit einem geringen Risiko für Komplikationen während der Geburt und einer von den Autorinnen und Autoren beschriebenen Standardbetreuung (spontane Wehentätigkeit und PDA) Wehenmittel während der Geburt. Damit sind diese Ergebnisse nur bedingt auf den deutschen Versorgungsalltag übertragbar, in welchem im Jahr 2017 26 % aller Gebärenden Wehenmittel während der Geburt erhielten [5]. Zudem stellt eine PDA in Deutschland keine Standardbetreuung dar und wurde im Jahr 2017 bei 23 % aller Gebärenden während der Geburt eingesetzt [5].

Auswahl der eingeschlossenen Studien

Alle eingeschlossenen Studien erfüllen die in Tabelle 2 festgelegten Einschlusskriterien. Ausgeschlossen wurden alle Studien, in denen die Einschlusskriterien dieses Evidenzberichts nicht bewertet werden konnten, die Analyseeinheit nicht die Gebärenden, sondern Kliniken waren und auf Grundlage der Beschreibungen in der Methodik davon ausgegangen wurde, dass keine prospektiv geplante Datenerfassung und Auswertung der erhobenen Daten erfolgte oder

ausschließlich Endpunkte berichtet wurden, welche den Endpunkten dieses Evidenzberichts nicht eindeutig zugeordnet werden konnten.

Analyseeinheit dieses Evidenzberichts sind die Gebärenden, weswegen ökologische Studien, in welchen ausschließlich eine Auswertung der Daten von Krankenhäusern ohne Berücksichtigung der Patientendaten der Gebärenden erfolgte [68,69], ausgeschlossen wurden.

1 systematische Übersicht [70] zu Interventionen, die Gynäkologinnen und Gynäkologen, Hebammen und Entbindungshelfer unterstützen sollen, den Anteil vaginaler Geburten nach einem Kaiserschnitt zu erhöhen, schloss neben der eingeschlossenen Studie (Lomas 1991) eine der aufgrund ihrer Analyseeinheit in diesem Evidenzbericht ausgeschlossene Studie zu Weiterbildungsmaßnahmen ein [69]. Diese von 1988 bis 1993 in den USA durchgeführte quasi-clusterrandomisierte Studie wurde mit dem Ziel initiiert, die Anzahl von vaginalen Geburten nach einem Kaiserschnitt über Weiterbildungsmaßnahmen in Form eines externen Reviewverfahrens zu erhöhen. Die Studie wurde aufgrund fehlender Strukturgleichheit der Interventionsgruppen nicht in diesen Evidenzbericht eingeschlossen. Für die Gruppe der Prüfintervention wurden Geburtskliniken mit geringen, mittleren und hohen Kaiserschnittraten zufällig ausgewählt, zusätzlich wurden weitere Geburtskliniken mit hohen Kaiserschnittraten der Prüfintervention zugewiesen. In der Gruppe mit der Vergleichsintervention ohne externes Reviewverfahren wurden alle nicht ausgewählten Geburtskliniken einschließlich der Geburtskliniken, welche die Intervention ablehnten, ausgewertet. Die Studie wurde als nicht randomisierte Interventionsstudie klassifiziert, eine Confounderadjustierung für individuelle Merkmale der Gebärenden erfolgte nicht. Daher lag die Kaiserschnitttrate zu Interventionsbeginn in den Geburtskliniken der Interventionsgruppe über der in der Vergleichsintervention. Die Auswertung basierte auf den Kaiserschnittraten von 45 Geburtskliniken aus der Gruppe der Prüfintervention und 120 Geburtskliniken der Vergleichsintervention. Es konnte weder für alle Geburtskliniken noch für Gruppen von Geburtskliniken mit unterschiedlichen Kaiserschnittraten ein signifikanter Unterschied zwischen der Prüf- und Vergleichsintervention nachgewiesen werden.

Alle Studien, in welchen Schulungs- oder Weiterbildungsmaßnahmen untersucht wurden, die die Einführung von Interventionen in überwiegend einkommensschwachen Ländern zur Verbesserung der Versorgung von Mutter und Neugeborenem untersuchten, die in Deutschland als Teil der Ausbildung des geburtshilflichen Personals oder als Standardmaßnahmen in der geburtshilflichen Versorgung bezeichnet werden können (z. B. Stimulierung des Neugeborenen nach der Geburt bei Anpassungsstörungen), wurden bereits auf Titel- und Abstractebene ausgeschlossen.

Eine prospektiv durchgeführte vergleichende Studie [67], in welcher die erhobenen Endpunkte nicht den in diesem Evidenzbericht interessierenden Endpunkten zugeordnet werden konnten, untersuchte den Einfluss des Geburtszeitpunktes, der Geburtenzahlen in den Geburtskliniken und der Qualifikation der die Geburt betreuenden Fachpersonen auf das Auftreten schwerwiegender und sehr schwerwiegender Ereignisse beim Neugeborenen. Die Studie wurde

in den Jahren 2004 bis 2015 in Österreich durchgeführt und schloss alle Geburten ab der vollendeten 34. SSW ein. Sehr schwerwiegende Ereignisse fassten als zusammengesetzter Endpunkt das Auftreten einer schweren metabolischen Acidose ($\text{pH} < 7,0$), eines geringen Apgar-Scores nach 5 Minuten ≤ 3 , die Notwendigkeit einer Herzdruckmassage / Wiederbelebung oder intrapartale beziehungsweise frühe neonatale Todesfälle zusammen. Schwerwiegende Ereignisse beschrieben das Auftreten einer fortgeschrittenen Acidose ($\text{pH} < 7,05$), eines Apgar-Scores nach 5 Minuten ≤ 6 , einer Verlegung auf die Neugeborenen-Intensivstation innerhalb von 24 Stunden nach der Geburt oder die Notwendigkeit einer Atemunterstützung. Da die in diesem kombinierten Endpunkt zusammengefassten Komponenten nicht von vergleichbarer klinischer Bedeutung waren, die Ergebnisse für diese Komponenten nicht einzeln berichtet wurden und die Variablen zur Erfahrung Teil der Variablenselektion waren, wurde auch diese Studie nicht in den Evidenzbericht aufgenommen. Die Auswertung basierte auf den Daten von 87 065 Neugeborenen. Sowohl sehr schwerwiegende als auch schwerwiegende Ereignisse traten häufiger bei Geburten auf, welche durch einen erfahrenen Oberarzt betreut wurden, als bei einer Betreuung durch einen Assistenzarzt (OR: 2,77; 95 %-KI: [2,77; 3,89] und OR: 1,74; 95 %-KI: [1,63; 1,86]). Zusätzlich wurde eine höhere Anzahl schwerwiegender Ereignisse in spezialisierten Perinatalzentren im Vergleich zu Geburtskliniken mit 500–1000 Geburten im Jahr nachgewiesen (OR: 2,03; 95 %-KI [1,37; 3,02]).

Es wurden mehrere Studien identifiziert, in denen das Ziel der Intervention die Verbesserung von Erfahrung / Expertise war und ausschließlich dieses und keiner der von der Leitliniengruppe vorgegebenen kritischen oder wichtigen Endpunkte erfasst und ausgewertet wurde. Diese Studien wurden bereits auf Titel- und Abstractebene ausgeschlossen. Eine aktuelle systematische Übersichtsarbeit [71] zur Wirksamkeit von Weiterbildungsmaßnahmen zur vaginalen Geburt von Feten in Beckenendlage konnte keine Studie identifizieren, welche die vorgegebenen Endpunkte untersuchte.

Zur Vermeidung von Selektionsbias war auch für diese Fragestellung primär der Einschluss von RCTs geplant. Da ausschließlich 2 RCT identifiziert werden konnten, in welchen der Einfluss einer durch Weiterbildungsmaßnahmen verbesserten Erfahrung / Expertise untersucht wird, wurden auch prospektiv geplante vergleichende Interventionsstudien mit zeitlich paralleler Kontrollgruppe und adäquater Confounderkontrolle zur Untersuchung von systemimmanenten Faktoren herangezogen.

Die in Tabelle 2 zusammengefassten Einschlusskriterien dieses Evidenzberichtes setzen eine prospektive Planung der einzuschließenden Studie voraus. Dies schließt eine eindeutige Beschreibung und prospektive Festlegung aller wesentlichen erklärenden Variablen ein, um eine selektive Berichterstattung und einen Publikationsbias zu vermeiden. Daher wurden explorative Studien, bei denen Ergebnisse multifaktorieller Modelle berichtet wurden, in welchen auch die Erfahrung / Expertise beschreibende Variablen Teil der Selektion waren, aufgrund ihres Studientyps ausgeschlossen [67,72-87]. In diesen Studien wurden die

untersuchten Einflussgrößen ab einem bestimmten Signifikanzniveau in das Modell eingeschlossen und nur für diese Variablen konnten Ergebnisse berichtet werden. Dieses Vorgehen führt zu einer Vorauswahl von Ergebnissen oberhalb eines bestimmten Signifikanzniveaus und einer daraus resultierenden Variablenselektion und einem resultierenden Publikationsbias.

Eine adäquate Confounderkontrolle sollte in nicht randomisierten Interventionsstudien sicherstellen, dass die beobachteten Unterschiede in den Endpunkten nicht auf systematische Unterschiede in der Zuteilung der Gebärenden in die Expositionsgruppen mit hoher oder geringerer Erfahrung / Expertise des Fachpersonals zurückgeführt werden können. Deshalb wurden ausschließlich Studien mit berichteten Ergebnissen aus Analysen mit adäquater Confounderkontrolle eingeschlossen. Diese Bedingung wurde endpunktspezifisch anhand der in den Ein- und Ausschlusskriterien definierten Auswahl der in den Studien ausgewerteten Gebärenden und der erfolgten Confounderadjustierung überprüft. Studien ohne entsprechende Adjustierung wurden aufgrund ihres Studientyps nach der Sichtung der Volltexte ausgeschlossen [66,67,78,82,88-115]).

10 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 Dammrisse 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 Endpunkt 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.

11 Literatur

1. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Die vaginale Geburt am Termin: angemeldetes Leitlinienvorhaben [online]. [Zugriff: 01.02.2017]. URL: <http://www.awmf.org/leitlinien/detail/anmeldung/1/II/015-083.html>.
2. Statistisches Bundesamt. Todesursachen bei Säuglingen: gestorbene Säuglinge nach Alter und ausgewählten Todesursachen der ICD-10 für 2015 [online]. [Zugriff: 13.11.2017]. URL: <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Gesundheit/Todesursachen/Tabellen/GestorbeneSaeuglinge.html>.
3. Bundesinstitut für Bevölkerungsforschung. Müttersterblichkeit [online]. [Zugriff: 13.11.2017]. URL: <http://www.bib-demografie.de/SharedDocs/Glossareintraege/DE/M/muettersterblichkeit.html>.
4. Zeitlin J, Mohangoo A, Delnord M. European perinatal health report: health and care of pregnant women and babies in Europe in 2010 [online]. [Zugriff: 09.12.2016]. URL: <http://www.europeristat.com/images/doc/Peristat%202013%20V2.pdf>.
5. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Geburtshilfe: Bundesauswertung zum Erfassungsjahr 2017; Qualitätsindikatoren [online]. 01.08.2018 [Zugriff: 24.08.2018]. URL: https://iqtig.org/downloads/auswertung/2017/16n1gebh/QSKH_16n1-GEBH_2017_BUAW_V02_2018-08-01.pdf.
6. Kolip P, Nolting HD, Zich K. Kaiserschnittgeburten in Deutschland: starker Anstieg und große regionale Unterschiede [online]. 2012 [Zugriff: 01.02.2017]. URL: https://faktencheck-gesundheit.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Themenblatt_VV_FCKaiserschnitt_2012.pdf.
7. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Klassifikation S3-Leitlinien [online]. [Zugriff: 19.10.2016]. URL: <http://www.awmf.org/leitlinien/awmf-regelwerk/II-entwicklung/awmf-regelwerk-01-planung-und-organisation/po-stufenklassifikation/klassifikation-s3.html>.
8. Dudenhausen JW. Praktische Geburtshilfe. Berlin: DeGruyter; 2011.
9. Krahl A, Schnepf W, Zu Sayn-Wittgenstein F. Die Bedeutung der Latenzphase der Geburt: eine historische Analyse. Z Geburtshilfe Neonatol 2016; 220(04): 155-165.
10. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Definitionen der Geburtsphasen: Rapid Report; Auftrag V16-01A. Köln: IQWiG; 2017. (IQWiG-Berichte; Band 517).
11. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G et al. GRADE guidelines; 2: framing the question and deciding on important outcomes. J Clin Epidemiol 2011; 64(4): 395-400.

12. Langer G, Meerpohl JJ, Perleth M, Gartlehner G, Kaminski-Hartenthaler A, Schunemann H. GRADE-Leitlinien; 2: Formulierung der Fragestellung und Entscheidung über wichtige Endpunkte. *Z Evid Fortbild Qual Gesundheitswes* 2012; 106(5): 369-376.
13. Schneider H, Husslein P, Schneifer KTM. *Die Geburtshilfe*. Heidelberg: Springer; 2006.
14. Diedrich K, Holzgreve W, Jonat W, Schultze-Mosgau A, Schneider KTM, Weiss JM (Ed). *Gynäkologie und Geburtshilfe*. Heidelberg: Springer; 2007.
15. Oxford Centre for Evidence-based Medicine. Levels of evidence [online]. In: *Essential Evidence Plus*. [Zugriff: 09.12.2016]. URL: https://www.essentialevidenceplus.com/product/ebm_loe.cfm?show=oxford.
16. Seo HJ, Kim SY, Lee YJ, Jang BH, Park JE, Sheen SS et al. A newly developed tool for classifying study designs in systematic reviews of interventions and exposures showed substantial reliability and validity. *J Clin Epidemiol* 2016; 70: 200-205.
17. Rose L, Schmied-Knittel I. Magie und Technik: moderne Geburt zwischen biografischem Event und kritischem Ereignis. In: Villa PI, Moebius S, Thiessen B (Ed). *Soziologie der Geburt: Diskurse, Praktiken und Perspektiven*. Frankfurt am Main: Campus-Verlag; 2011. S. 75-100.
18. 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.
19. 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.
20. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
21. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J et al. GRADE guidelines; 3: rating the quality of evidence. *J Clin Epidemiol* 2011; 64(4): 401-406.
22. Meerpohl JJ, Langer G, Perleth M, Gartlehner G, Kaminski-Hartenthaler A, Schunemann H. GRADE-Leitlinien; 3: Bewertung der Qualität der Evidenz (Vertrauen in die Effektschätzer). *Z Evid Fortbild Qual Gesundheitswes* 2012; 106(6): 449-456.
23. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P et al. GRADE guidelines; 4: rating the quality of evidence; study limitations (risk of bias). *J Clin Epidemiol* 2011; 64(4): 407-415.
24. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009; 6(5): e1000065.
25. Centre for Evidence-Based Medicine. OCEBM levels of evidence [online]. [Zugriff: 26.04.2017]. URL: <http://www.cebm.net/ocebmllevels-of-evidence/>.

26. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R et al. GRADE guidelines; 12: preparing summary of findings tables; binary outcomes. *J Clin Epidemiol* 2013; 66(2): 158-172.
27. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
28. 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.
29. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 5.0. Köln: IQWiG; 2017. URL: <https://www.iqwig.de/download/Allgemeine-Methoden-Version-5-0.pdf>.
30. 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.
31. Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008.
32. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.
33. Nussbaumer B, Gartlehner G, Kien C, Kaminski-Hartenthaler A, Langer G, Meerpohl JJ et al. Grade Leitlinien; 15: von der Evidenz zur Empfehlung; Determinanten, die Richtung und Stärke einer Empfehlung bestimmen. *Z Evid Fortbild Qual Gesundhwes* 2014; 108(7): 421-431.
34. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P et al. GRADE guidelines; 11: making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol* 2013; 66(2): 151-157.
35. Kaminski-Hartenthaler A, Gartlehner G, Kien C, Meerpohl JJ, Langer G, Perleth M et al. GRADE-Leitlinien; 11: Gesamtbeurteilung des Vertrauens in Effektschätzer für einen einzelnen Studienendpunkt und für alle Endpunkte. *Z Evid Fortbild Qual Gesundhwes* 2013; 107(9-10): 638-645.
36. Meerpohl JJ, Langer G, Perleth M, Gartlehner G, Kaminski-Hartenthaler A, Schunemann H. GRADE-Leitlinien; 4: Bewertung der Qualität der Evidenz; Studienlimitationen (Risiko für Bias). *Z Evid Fortbild Qual Gesundhwes* 2012; 106(6): 457-469.
37. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J et al. GRADE guidelines; 5: rating the quality of evidence; publication bias. *J Clin Epidemiol* 2011; 64(12): 1277-1282.
38. Nolting A, Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A et al. GRADE Leitlinien; 5: Einschätzung der Qualität der Evidenz; Publikationsbias. *Z Evid Fortbild Qual Gesundhwes* 2012; 106(9): 670-676.

39. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D et al. GRADE guidelines; 6: rating the quality of evidence; imprecision. *J Clin Epidemiol* 2011; 64(12): 1283-1293.
40. Kulig M, Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A et al. GRADE Leitlinien; 6: Einschätzung der Qualität der Evidenz; unzureichende Präzision. *Z Evid Fortbild Qual Gesundheitswes* 2012; 106(9): 677-688.
41. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. GRADE guidelines; 7: rating the quality of evidence; inconsistency. *J Clin Epidemiol* 2011; 64(12): 1294-1302.
42. Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A, Schunemann HJ. GRADE Leitlinien; 7: Einschätzung der Qualität der Evidenz; Inkonsistenz. *Z Evid Fortbild Qual Gesundheitswes* 2012; 106(10): 733-744.
43. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. GRADE guidelines; 8: rating the quality of evidence; indirectness. *J Clin Epidemiol* 2011; 64(12): 1303-1310.
44. Rasch A, Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A et al. GRADE Leitlinien; 8: Einschätzung der Qualität der Evidenz; Indirektheit. *Z Evid Fortbild Qual Gesundheitswes* 2012; 106(10): 745-753.
45. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P et al. GRADE guidelines; 9: rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12): 1311-1316.
46. Kien C, Gartlehner G, Kaminski-Hartenthaler A, Meerpohl JJ, Flamm M, Langer G et al. GRADE-Leitlinien; 9: Heraufstufen der Qualität der Evidenz. *Z Evid Fortbild Qual Gesundheitswes* 2013; 107(3): 249-255.
47. Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA et al. GRADE guidelines; 13: preparing summary of findings tables and evidence profiles; continuous outcomes. *J Clin Epidemiol* 2013; 66(2): 173-183.
48. Langer G, Meerpohl JJ, Perleth M, Gartlehner G, Schunemann H. GRADE-Leitlinien; 12: Erstellen von „summary-of-findings“-Tabellen; dichotome Endpunkte. *Z Evid Fortbild Qual Gesundheitswes* 2013; 107(9-10): 646-664.
49. Schell LK, Meerpohl JJ, Gartlehner G, Langer G, Perleth M, Schunemann HJ. GRADE Leitlinien; 13: Erstellen von summary-of-findings-Tabellen und Evidenzprofilen; kontinuierliche Endpunkte. *Z Evid Fortbild Qual Gesundheitswes* 2014; 108(5-6): 333-347.
50. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, Neumann I, Mustafa RA, Mbuagbaw L et al. Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. *J Clin Epidemiol* 2016; 74: 7-18.

51. Santesso N, Carrasco-Labra A, Langendam M, Brignardello-Petersen R, Mustafa RA, Heus P et al. Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. *J Clin Epidemiol* 2016; 74: 28-39.
52. Belghiti J, Coulm B, Kayem G, Blondel B, Deneux-Tharaux C. Oxytocin administration during labor: results from the 2010 French national perinatal survey [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2013; 42(7): 662-670.
53. Chaillet N, Dumont A, Abrahamowicz M, Pasquier JC, Audibert F, Monnier P et al. A cluster-randomized trial to reduce cesarean delivery rates in Quebec. *N Engl J Med* 2015; 372(18): 1710-1721.
54. Chuilon AL, Le Ray C, Prunet C, Blondel B. Episiotomy in France in 2010: variations according to obstetrical context and place of birth [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2016; 45(7): 691-700.
55. Coulm B, Le Ray C, Lelong N, Drewniak N, Zeitlin J, Blondel B. Obstetric interventions for low-risk pregnant women in France: do maternity unit characteristics make a difference? *Birth* 2012; 39(3): 183-191.
56. Finnstrom O, Berg G, Norman A, Otterblad Olausson P. Size of delivery unit and neonatal outcome in Sweden: a catchment area analysis. *Acta Obstet Gynecol Scand* 2006; 85(1): 63-67.
57. Karalis E, Gissler M, Tapper AM, Ulander VM. Effect of hospital size and on-call arrangements on intrapartum and early neonatal mortality among low-risk newborns in Finland. *Eur J Obstet Gynecol Reprod Biol* 2016; 198: 116-119.
58. Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs audit and feedback to implement practice guidelines: delivery after previous cesarean section. *J Am Med Assoc* 1991; 265(17): 2202-2207.
59. Wehberg S, Guldberg R, Gradel KO, Kesmodel US, Munk L, Andersson CB et al. Risk factors and between-hospital variation of caesarean section in Denmark: a cohort study. *BMJ Open* 2018; 8(2): e019120.
60. Royal College of Obstetricians and Gynaecologists. The OASI care bundle: a quality improvement project to reduce rates of severe perineal tears that occur during childbirth [online]. In: ISRCTN Registry. 19.03.2018 [Zugriff: 03.11.2018]. URL: <http://isrctn.com/ISRCTN12143325>.
61. Milupa Nutricia. Geburtenliste Deutschland 2017: für medizinische Fachkreise [online]. 01.2018 [Zugriff: 05.09.2018]. URL: <https://www.presseportal.de/download/document/464725-milupa-geburtenliste-2017.pdf>.
62. Wörterbuch der deutschen Gegenwartssprache (1964–1977), kuratiert und bereitgestellt durch das Digitale Wörterbuch der deutschen Sprache. "Berufserfahrung" [online]. [Zugriff: 29.06.2018]. URL: <https://www.dwds.de/wb/wdg/Berufserfahrung>.

63. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen zur Qualitätssicherung der Versorgung von Früh- und Reifgeborenen gemäß § 136 Absatz 1 Nummer 2 SGB V in Verbindung mit § 92 Abs. 1 Satz 2 Nr. 13 SGB V: (Qualitätssicherungs-Richtlinie Früh- und Reifgeborene/QFR-RL) [online]. 17.05.2018 [Zugriff: 05.09.2018]. URL: https://www.g-ba.de/downloads/62-492-1646/QFR-RL_2018-05-17_iK-2018-08-25.pdf.
64. Statistisches Bundesamt. Anteil der verschiedenen Entbindungsarten an den Geburten 1994 bis 2013 [online]. [Zugriff: 24.01.2017]. URL: http://www.gbe-bund.de/gbe10/ergebnisse.prc_fid_anzeige?p_fid=24246&p_fund_typ=TAB&p_sprachkz=D&p_uid=gast&p_aid=35573339&p_prot=2.
65. Chen I, Opiyo N, Tavender E, Mortazhejri S, Rader T, Petkovic J et al. Non-clinical interventions for reducing unnecessary caesarean section. *Cochrane Database Syst Rev* 2018; (9): CD005528.
66. Tracy SK, Sullivan E, Dahlen H, Black D, Wang YA, Tracy MB. Does size matter? A population-based study of birth in lower volume maternity hospitals for low risk women. *BJOG* 2006; 113(1): 86-96.
67. Reif P, Pichler G, Griesbacher A, Lehner G, Schöll W, Lang U et al. Do time of birth, unit volume, and staff seniority affect neonatal outcome in deliveries at $\geq 34^{+0}$ weeks of gestation? *BJOG* 06.12.2017 [Epub ahead of print].
68. Moster D, Lie RT, Markestad T. Neonatal mortality rates in communities with small maternity units compared with those having larger maternity units. *BJOG* 2001; 108(9): 904-909.
69. Bickell NA, Zdeb MS, Applegate MS, Roohan PJ, Siu AL. Effect of external peer review on cesarean delivery rates: a statewide program. *Obstet Gynecol* 1996; 87(5): 664-667.
70. Lundgren I, Smith V, Nilsson C, Vehvilainen-Julkunen K, Nicoletti J, Devane D et al. Clinician-centred interventions to increase vaginal birth after caesarean section (VBAC): a systematic review. *BMC Pregnancy Childbirth* 2015; 15(1): 16.
71. Walker S, Breslin E, Scamell M, Parker P. Effectiveness of vaginal breech birth training strategies: an integrative review of the literature. *Birth* 2017; 44(2): 101-109.
72. Acosta CD, Knight M, Lee HC, Kurinczuk JJ, Gould JB, Lyndon A. The continuum of maternal sepsis severity: incidence and risk factors in a population-based cohort study. *PLoS One* 2013; 8(7): e67175.
73. Aiken AR, Aiken CE, Alberry MS, Brockelsby JC, Scott JG. Management of fetal malposition in the second stage of labor: a propensity score analysis. *Am J Obstet Gynecol* 2015; 212(3): 355.e1-355.e7.
74. Azria E, Le Meaux JP, Khoshnood B, Alexander S, Subtil D, Goffinet F. Factors associated with adverse perinatal outcomes for term breech fetuses with planned vaginal delivery. *Am J Obstet Gynecol* 2012; 207(4): 285.e1-285.e9.

75. Burns LR, Geller SE, Wholey DR. The effect of physician factors on the cesarean section decision. *Med Care* 1995; 33(4): 365-382.
76. Goldman G, Pineault R, Bilodeau H, Blais R. Effects of patient, physician and hospital characteristics on the likelihood of vaginal birth after previous cesarean section in Quebec. *CMAJ* 1990; 143(10): 1017-1024.
77. Joyce R, Webb R, Peacock J. Predictors of obstetric intervention rates: case-mix, staffing levels and organisational factors of hospital of birth. *J Obstet Gynaecol* 2002; 22(6): 618-625.
78. Joyce R, Webb R, Peacock JL. Associations between perinatal interventions and hospital stillbirth rates and neonatal mortality. *Arch Dis Child Fetal Neonatal Ed* 2004; 89(1): F51-F56.
79. Klein MC, Spence A, Kaczorowski J, Kelly A, Grzybowski S. Does delivery volume of family physicians predict maternal and newborn outcome? *CMAJ* 2002; 166(10): 1257-1263.
80. Le Ray C, Carayol M, Zeitlin J, Breart G, Goffinet F. Level of perinatal care of the maternity unit and rate of cesarean in low-risk nulliparas. *Obstet Gynecol* 2006; 107(6): 1269-1277.
81. Le Ray C, Zeitlin J, Jarreau PH, Breart G, Goffinet F. The influence of level of care on admission to neonatal care for babies of low-risk nullipara. *Eur J Obstet Gynecol Reprod Biol* 2009; 144(1): 21-26.
82. Luthy DA, Malmgren JA, Zingheim RW, Leininger CJ. Physician contribution to a cesarean delivery risk model. *Am J Obstet Gynecol* 2003; 188(6): 1579-1585.
83. McClelland S, Gorfinkle N, Arslan AA, Benedetto-Anzai MT, Cheon T, Anzai Y. Factors associated with cesarean delivery rates: a single-institution experience. *Matern Health Neonatol Perinatol* 2017; 3: 8.
84. Roman H, Carayol M, Watier L, Le Ray C, Breart G, Goffinet F. Planned vaginal delivery of fetuses in breech presentation at term: prenatal determinants predictive of elevated risk of cesarean delivery during labor. *Eur J Obstet Gynecol Reprod Biol* 2008; 138(1): 14-22.
85. Sagot P, Gouyon-Cornet B, Gouyon JB. Cesarean section in the Burgundy perinatal network: reasons for differences in the management of mothers [Französisch]. *Gynecol Obstet Fertil* 2003; 31(7-8): 606-613.
86. Sandall J, Murrells T, Dodwell M, Gibson R, Bewley S, Coxon K et al. The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study. *Health Serv Deliv Res* 2014; 2(38).
87. Vankan E, Schoorel EN, Van Kuijk SM, Mol BJ, Nijhuis JG, Aardenburg R et al. Practice variation of vaginal birth after cesarean and the influence of risk factors at patient level: a retrospective cohort study. *Acta Obstet Gynecol Scand* 2017; 96(2): 158-165.

88. Allen J, Hippisley-Cox J, Pringle M, Groom L. Assisted delivery in the teenage population: the effect of inter-hospital variation, deprivation, and age. *Int J Adolesc Med Health* 2003; 15(4): 341-347.
89. Chang ALS, Noah MS, Laros RK Jr. Obstetric attending physician characteristics and their impact on vacuum and forceps delivery rates: University of California at San Francisco experience from 1977 to 1999. *Am J Obstet Gynecol* 2002; 186(6): 1299-1303.
90. DeMott RK, Sandmire HF. The Green Bay cesarean section study; I: the physician factor as a determinant of cesarean birth rates. *Am J Obstet Gynecol* 1990; 162(6): 1593-1599.
91. DeMott RK, Sandmire HF. The Green Bay cesarean section study; II: the physician factor as a determinant of cesarean birth rates for failed labor. *Am J Obstet Gynecol* 1992; 166(6 Pt 1): 1799-1806.
92. Escuriet R, Pueyo M, Biescas H, Colls C, Espiga I, White J et al. Obstetric interventions in two groups of hospitals in Catalonia: a cross-sectional study. *BMC Pregnancy Childbirth* 2014; 14: 143.
93. Escuriet R, Pueyo MJ, Perez-Botella M, Espada X, Salgado I, Gomez A et al. Cross-sectional study comparing public and private hospitals in Catalonia: is the practice of routine episiotomy changing? *BMC Health Serv Res* 2015; 15: 95.
94. Goode KT, Weiss PM, Koller C, Kimmel S, Hess LW. Episiotomy rates in private vs. resident service deliveries: a comparison. *J Reprod Med* 2006; 51(3): 190-192.
95. Gossett DR, Dunsmoor Su R. Episiotomy practice in a community hospital setting. *J Reprod Med* 2008; 53(10): 803-808.
96. Griffiths AN, Avasarala S, Wiener JJ. A prospective observational study of emergency caesarean section rates and the effect of the labour ward experience. *J Obstet Gynaecol* 2005; 25(7): 666-668.
97. Konstantiniuk P, Kern I, Giuliani A, Kainer F. The midwife factor in obstetric procedures and neonatal outcome. *J Perinat Med* 2002; 30(3): 242-249.
98. Kozhimannil KB, Thao V, Peiyin H, Tilden E, Caughey AB, Snowden JM. Association between hospital birth volume and maternal morbidity among low-risk pregnancies in rural, urban, and teaching hospitals in the United States. *Am J Perinatol* 2016; 33(6): 590-599.
99. Kyser KL, Lu X, Santillan D, Santillan M, Caughey AB, Wilson MC et al. Forceps delivery volumes in teaching and nonteaching hospitals: are volumes sufficient for physicians to acquire and maintain competence? *Acad Med* 2014; 89(1): 71-76.
100. Lin HC, Xirasagar S. Institutional factors in cesarean delivery rates: policy and research implications. *Obstet Gynecol* 2004; 103(1): 128-136.
101. Lundsberg LS, Illuzzi JL, Garipey AM, Sheth SS, Pettker CM, Lee HC et al. Variation in hospital intrapartum practices and association with cesarean rate. *J Obstet Gynecol Neonatal Nurs* 2017; 46(1): 5-17.

102. Maeda E, Ishihara O, Tomio J, Sato A, Terada Y, Kobayashi Y et al. Cesarean section rates and local resources for perinatal care in Japan: a nationwide ecological study using the national database of health insurance claims. *J Obstet Gynaecol Res* 2018; 44(2): 208-216.
103. Mangin M, Ramanah R, Aouar Z, Courtois L, Collin A, Cossa S et al. Operative delivery data in France for 2007: results of a national survey within teaching hospitals [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2010; 39(2): 121-132.
104. Parazzini F, Cipriani S, Bulfoni G, Bulfoni C, Bellu R, Zanini R et al. Mode of delivery and level of neonatal care in Lombardy: a descriptive analysis according to volume of care. *Ital J Pediatr* 2015; 41: 24.
105. Sandmire HF, DeMott RK. The Green Bay cesarean section study; III: falling cesarean birth rates without a formal curtailment program. *Am J Obstet Gynecol* 1994; 170(6): 1790-1799.
106. Sandmire HF, DeMott RK. The Green Bay cesarean section study; IV: the physician factor as a determinant of cesarean birth rates for the large fetus. *Am J Obstet Gynecol* 1996; 174(5): 1557-1564.
107. Simoes E, Kunz S, Bosing-Schwenkglenks M, Schmahl FW. Mütterliche Anämie post partum: Entwicklungstendenzen und Variabilität unter Berücksichtigung unterschiedlicher Klinikategorien; Untersuchung auf der Basis der Perinatalerhebung in Baden-Württemberg. *Z Geburtshilfe Neonatol* 2004; 208(5): 184-189.
108. Takacs L, Seidlerova J. Psychosocial climate in maternity hospitals from the perspective of parturients I: results from a national survey on perinatal care satisfaction using a representative sample of 1195 Czech parturients [Tschechisch]. *Ceska Gynekol* 2013; 78(2): 157-168.
109. Vendittelli F, Barasinski C, Pereira B, Lemery D. Incidence of immediate postpartum hemorrhages in French maternity units: a prospective observational study (HERA study). *BMC Pregnancy Childbirth* 2016; 16: 242.
110. Heller G, Schnell R, Richardson DK, Misselwitz B, Schmidt S. Hat die Größe der Geburtsklinik Einfluss auf das neonatale Überleben? Schätzung von „vermeidbaren“ Todesfällen in Hessen 1990–2000. *Dtsch Med Wochenschr* 2003; 128(13): 657-662.
111. LeFevre M, Sanner L, Anderson S, Tsutakawa R. The relationship between neonatal mortality and hospital level. *J Fam Pract* 1992; 35(3): 259-264.
112. Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. *Am J Public Health* 1990; 80(7): 819-823.
113. Pyykönen A, Gissler M, Jakobsson M, Petäjä J, Tapper AM. Determining obstetric patient safety indicators: the differences in neonatal outcome measures between different-sized delivery units. *BJOG* 2014; 121(4): 430-437.

114. Viisainen K, Gissler M, Hemminki E. Birth outcomes by level of obstetric care in Finland: a catchment area based analysis. *J Epidemiol Community Health* 1994; 48(4): 400-405.
115. Pyykönen A, Gissler M, Jakobsson M, Lehtonen L, Tapper AM. The rate of obstetric anal sphincter injuries in Finnish obstetric units as a patient safety indicator. *Eur J Obstet Gynecol Reprod Biol* 2013; 169(1): 33-38.
116. Wong SSL, 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.
117. Lefebvre C, Manheimer E, Glanville J. Searching for studies [online]. In: Higgings JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions: version 5.1.0*. 03.2011 [Zugriff: 17.02.2017]. URL: http://handbook.cochrane.org/chapter_6/6_searching_for_studies.htm.

12 Studienlisten

12.1 Liste der eingeschlossenen Studien

Belghiti 2013

Belghiti J, Coulm B, Kayem G, Blondel B, Deneux-Tharoux C. Oxytocin administration during labor: results from the 2010 French national perinatal survey [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2013; 42(7): 662-670.

Chailet 2015

Chaillet N, Dumont A, Abrahamowicz M, Pasquier JC, Audibert F, Monnier P et al. A cluster-randomized trial to reduce cesarean delivery rates in Quebec. *N Engl J Med* 2015; 372(18): 1710-1721.

Chuilon 2016

Chuilon AL, Le Ray C, Prunet C, Blondel B. Episiotomy in France in 2010: variations according to obstetrical context and place of birth [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2016; 45(7): 691-700.

Coulm 2012

Coulm B, Le Ray C, Lelong N, Drewniak N, Zeitlin J, Blondel B. Obstetric interventions for low-risk pregnant women in France: do maternity unit characteristics make a difference? *Birth* 2012; 39(3): 183-191.

Finnström 2006

Finnström O, Berg G, Norman A, Otterblad Olausson P. Size of delivery unit and neonatal outcome in Sweden: a catchment area analysis. *Acta Obstet Gynecol Scand* 2006; 85(1): 63-67.

Karalis 2016

Karalis E, Gissler M, Tapper AM, Ulander VM. Effect of hospital size and on-call arrangements on intrapartum and early neonatal mortality among low-risk newborns in Finland. *Eur J Obstet Gynecol Reprod Biol* 2016; 198: 116-119.

Lomas 1991

Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs audit and feedback to implement practice guidelines: delivery after previous cesarean section. *J Am Med Assoc* 1991; 265(17): 2202-2207.

Wehberg 2018

Wehberg S, Guldberg R, Gradel KO, Kesmodel US, Munk L, Andersson CB et al. Risk factors and between-hospital variation of caesarean section in Denmark: a cohort study. *BMJ Open* 2018; 8(2): e019120.

12.2 Liste der gesichteten systematischen Übersichten

1. Catling-Paull C, Johnston R, Ryan C, Foureur MJ, Homer CSE. Non-clinical interventions that increase the uptake and success of vaginal birth after caesarean section: a systematic review. *J Adv Nurs* 2011; 67(8): 1662-1676.
2. Chaillet N, Dumont A. Evidence-based strategies for reducing cesarean section rates: a meta-analysis. *Birth* 2007; 34(1): 53-64.
3. Chang YS, Coxon K, Portela AG, Furuta M, Bick D. Interventions to support effective communication between maternity care staff and women in labour: a mixed-methods systematic review. *Midwifery* 2018; 59: 4-16.
4. Chen I, Opiyo N, Tavender E, Mortazhejri S, Rader T, Petkovic J et al. Non-clinical interventions for reducing unnecessary caesarean section. *Cochrane Database Syst Rev* 2018; (9): CD005528.
5. Gijbels H, O'Connell R, Dalton-O'Connor C, O'Donovan M. A systematic review evaluating the impact of post-registration nursing and midwifery education on practice. *Nurse Educ Pract* 2010; 10(2): 64-69.
6. Lundgren I, Smith V, Nilsson C, Vehvilainen-Julkunen K, Nicoletti J, Devane D et al. Clinician-centred interventions to increase vaginal birth after caesarean section (VBAC): a systematic review. *BMC Pregnancy Childbirth* 2015; 15(1): 16.
7. Merien AER, Van de Ven J, Mol BW, Houterman S, Oei SG. Multidisciplinary team training in a simulation setting for acute obstetric emergencies: a systematic review. *Obstet Gynecol* 2010; 115(5): 1021-1031.
8. Neogi SB, Malhotra S, Zodpey S, Mohan P. Does facility based newborn care improve neonatal outcomes? A review of evidence. *Indian Pediatr* 2012; 49(8): 651-658.
9. Walker S, Breslin E, Scamell M, Parker P. Effectiveness of vaginal breech birth training strategies: an integrative review of the literature. *Birth* 2017; 44(2): 101-109.

12.3 Liste der ausgeschlossenen Publikationen zum Thema mit Ausschlussgründen

Nicht E1 (Population)

1. Di Lallo D, Perucci CA, Bertollini R, Mallone S. Cesarean section rates by type of maternity unit and level of obstetric care: an area-based study in central Italy. *Prev Med* 1996; 25(2): 178-185.
2. Elferink-Stinkens PM, Brand R, Amelink-Verburg MP, Merkus JMWM, Den Ouden AL, Van Hemel OJS. Randomised clinical trial on the effect of the Dutch obstetric peer review system. *Eur J Obstet Gynecol Reprod Biol* 2002; 102(1): 21-30.
3. Hemminki E, Gissler M. Variation in obstetric care within and between hospital levels in Finland. *Br J Obstet Gynaecol* 1994; 101(10): 851-857.

4. Hueston WJ, Lewis-Stevenson S. Provider distribution and variations in statewide cesarean section rates. *J Community Health* 2001; 26(1): 1-10.
5. Kim YM, Kim SY. Factors on the gap between predicted cesarean section rate and real cesarean section rate in tertiary hospitals. *Korean J Women Health Nurs* 2012; 18(3): 200-208.
6. Moster D, Lie RT, Markestad T. Neonatal mortality rates in communities with small maternity units compared with those having larger maternity units. *BJOG* 2001; 108(9): 904-909.
7. Rowe RE, Townend J, Brocklehurst P, Knight M, Macfarlane A, McCourt C et al. Service configuration, unit characteristics and variation in intervention rates in a national sample of obstetric units in England: an exploratory analysis. *BMJ Open* 2014; 4(5): e005551.

Nicht E2 (Prüfintervention)

1. Allanson ER, Tuncalp Ö, Vogel JP, Khan DN, Oladapo OT, Long Q et al. Implementation of effective practices in health facilities: a systematic review of cluster randomised trials. *BMJ Glob Health* 2017; 2(2): e000266.
2. Almeida LM, Caldas J, Ayres-de-Campos D, Salcedo-Barrientos D, Dias S. Maternal healthcare in migrants: a systematic review. *Matern Child Health J* 2013; 17(8): 1346-1354.
3. Cromi A, Bonzini M, Uccella S, Serati M, Bogani G, Pozzo N et al. Provider contribution to an episiotomy risk model. *J Matern Fetal Neonatal Med* 2015; 28(18): 2201-2206.
4. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: a systematic review of the effect of continuing medical education strategies. *JAMA* 1995; 274(9): 700-705.
5. Dunphy BC, Cantwell R, Bourke S, Fleming M, Smith B, Joseph KS et al. Cognitive elements in clinical decision-making: toward a cognitive model for medical education and understanding clinical reasoning. *Adv in Health Sci Educ* 2010; 15(2): 229-250.
6. Edmonds JK, O'Hara M, Clarke SP, Shah NT. Variation in cesarean birth rates by labor and delivery nurses. *J Obstet Gynecol Neonatal Nurs* 2017; 46(4): 486-493.
7. Fritz J, Walker DM, Cohen S, Angeles G, Lamadrid-Figueroa H. Can a simulation-based training program impact the use of evidence based routine practices at birth? Results of a hospital-based cluster randomized trial in Mexico. *PLoS One* 2017; 12(3): e0172623.
8. Karalis E, Gissler M, Tapper AM, Ulander VM. Influence of time of delivery on risk of adverse neonatal outcome in different size of delivery units: a retrospective cohort study in Finland. *J Matern Fetal Neonatal Med* 20.12.2017 [Epub ahead of print].
9. Lassi ZS, Cometto G, Huicho L, Bhutta ZA. Quality of care provided by mid-level health workers: systematic review and meta-analysis. *Bull World Health Organ* 2013; 91(11): 824-833I.

10. Manohar S, Woods CFG, Lindow SW. Individual consultant practice does not affect the overall intervention rate: a 6-year study. *J Perinat Med* 2015; 43(1): 37-41.
11. Moster D, Lie RT, Markestad T. Relation between size of delivery unit and neonatal death in low risk deliveries: population based study. *Arch Dis Child Fetal Neonatal Ed* 1999; 80(3): F221-F225.
12. Räisänen S, Vehviläinen-Julkunen K, Heinonen S. Need for and consequences of episiotomy in vaginal birth: a critical approach. *Midwifery* 2010; 26(3): 348-356.
13. Signorelli C, Cattaruzza MS, Osborn JF. Risk factors for caesarean section in Italy: results of a multicentre study. *Public Health* 1995; 109(3): 191-199.
14. Walker D, Cohen S, Fritz J, Olvera M, Lamadrid-Figueroa H, Cowan JG et al. Team training in obstetric and neonatal emergencies using highly realistic simulation in Mexico: impact on process indicators. *BMC Pregnancy Childbirth* 2014; 14: 367.
15. Walker DM, Cohen SR, Fritz J, Olvera-Garcia M, Zelek ST, Fahey JO et al. Impact evaluation of PRONTO Mexico: a simulation-based program in obstetric and neonatal emergencies and team training. *Simul Healthc* 2016; 11(1): 1-9.
16. Webb DA, Culhane J. Hospital variation in episiotomy use and the risk of perineal trauma during childbirth. *Birth* 2002; 29(2): 132-136.
17. Yee LM, Liu LY, Grobman WA. The relationship between obstetricians' cognitive and affective traits and their patients' delivery outcomes. *Am J Obstet Gynecol* 2014; 211(6): 692.e1-692.e6.

Nicht E3 (Vergleichsintervention)

1. Ott J, Gritsch E, Pils S, Kratschmar S, Promberger R, Seemann R et al. A retrospective study on perineal lacerations in vaginal delivery and the individual performance of experienced midwives. *BMC Pregnancy Childbirth* 2015; 15: 270.
2. Riddell CA, Hutcheon JA, Strumpf EC, Abenhaim HA, Kaufman JS. Inter-institutional variation in use of caesarean delivery for labour dystocia. *J Obstet Gynaecol Can* 2017; 39(11): 988-995.

Nicht E4 (Endpunkte)

1. Althabe F, Buekens P, Bergel E, Belizan JM, Campbell MK, Moss N et al. A behavioral intervention to improve obstetrical care. *N Engl J Med* 2008; 358(18): 1929-1940.
2. Deneux-Tharoux C, Dupont C, Colin C, Rabilloud M, Touzet S, Lansac J et al. Multifaceted intervention to decrease the rate of severe postpartum haemorrhage: the PITHAGORE6 cluster-randomised controlled trial. *BJOG* 2010; 117(10): 1278-1287.
3. Epstein AJ, Srinivas SK, Nicholson S, Herrin J, Asch DA. Association between physicians' experience after training and maternal obstetrical outcomes: cohort study. *BMJ* 2013; 346: f1596.

4. Hehir MP, Ananth CV, Wright JD, Siddiq Z, D'Alton ME, Friedman AM. Severe maternal morbidity and comorbid risk in hospitals performing <1000 deliveries per year. *Am J Obstet Gynecol* 2017; 216(2): 179.e1-179.e12.
5. Hemminki E, Heino A, Gissler M. Should births be centralised in higher level hospitals? Experiences from regionalised health care in Finland. *BJOG* 2011; 118(10): 1186-1195.
6. Merriam AA, Wright JD, Siddiq Z, D'Alton ME, Friedman AM, Ananth CV et al. Risk for postpartum hemorrhage, transfusion, and hemorrhage-related morbidity at low, moderate, and high volume hospitals. *J Matern Fetal Neonatal Med* 2018; 31(8): 1025-1034.
7. Reif P, Pichler G, Griesbacher A, Lehner G, Schöll W, Lang U et al. Do time of birth, unit volume, and staff seniority affect neonatal outcome in deliveries at $\geq 34^{+0}$ weeks of gestation? *BJOG* 06.12.2017 [Epub ahead of print].

Nicht E5 (Studenttyp)

1. Acosta CD, Knight M, Lee HC, Kurinczuk JJ, Gould JB, Lyndon A. The continuum of maternal sepsis severity: incidence and risk factors in a population-based cohort study. *PLoS One* 2013; 8(7): e67175.
2. Aiken AR, Aiken CE, Alberry MS, Brockelsby JC, Scott JG. Management of fetal malposition in the second stage of labor: a propensity score analysis. *Am J Obstet Gynecol* 2015; 212(3): 355.e1-355.e7.
3. Aiken CE, Aiken AR, Brockelsby JC, Scott JG. Factors influencing the likelihood of instrumental delivery success. *Obstet Gynecol* 2014; 123(4): 796-803.
4. Aiken CE, Aiken AR, Park H, Brockelsby JC, Prentice A. Factors associated with adverse clinical outcomes among obstetrics trainees. *Med Educ* 2015; 49(7): 674-683.
5. Allen J, Hippisley-Cox J, Pringle M, Groom L. Assisted delivery in the teenage population: the effect of inter-hospital variation, deprivation, and age. *Int J Adolesc Med Health* 2003; 15(4): 341-347.
6. Azria E, Le Meaux JP, Khoshnood B, Alexander S, Subtil D, Goffinet F. Factors associated with adverse perinatal outcomes for term breech fetuses with planned vaginal delivery. *Am J Obstet Gynecol* 2012; 207(4): 285.e1-285.e9.
7. Biro MA, Knight M, Wallace E, Papacostas K, East C. Is place of birth associated with mode of birth? The effect of hospital on caesarean section rates in a public metropolitan health service. *Aust N Z J Obstet Gynaecol* 2014; 54(1): 64-70.
8. Boucoiran I, Bafghi A, Delotte J, Valerio L, Bongain A. Learning use of spatulas: risks of perineal injuries and newborns morbidity [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2010; 39(3): 224-230.
9. Burns LR, Geller SE, Wholey DR. The effect of physician factors on the cesarean section decision. *Med Care* 1995; 33(4): 365-382.

10. Calvert KL, McGurgan PM, Debenham EM, Gratwick FJ, Maouris P. Emergency obstetric simulation training: how do we know where we are going, if we don't know where we have been? *Aust N Z J Obstet Gynaecol* 2013; 53(6): 509-516.
11. Chang ALS, Noah MS, Laros RK Jr. Obstetric attending physician characteristics and their impact on vacuum and forceps delivery rates: University of California at San Francisco experience from 1977 to 1999. *Am J Obstet Gynecol* 2002; 186(6): 1299-1303.
12. Clapp MA, James KE, Melamed A, Ecker JL, Kaimal AJ. Hospital volume and cesarean delivery among low-risk women in a nationwide sample. *J Perinatol* 2018; 38(2): 127-131.
13. Clapp MA, Melamed A, Robinson JN, Shah N, Little SE. Obstetrician volume as a potentially modifiable risk factor for cesarean delivery. *Obstet Gynecol* 2014; 124(4): 697-703.
14. Clark SL, Xu W, Porter TF, Love D. Institutional influences on the primary cesarean section rate in Utah, 1992 to 1995. *Am J Obstet Gynecol* 1998; 179(4): 841-845.
15. Das JK, Kumar R, Salam RA, Lassi ZS, Bhutta ZA. Evidence from facility level inputs to improve quality of care for maternal and newborn health: interventions and findings. *Reprod Health* 2014; 11(Suppl 2): S4.
16. David S, Mamelle N, Riviere O. Estimation of an expected caesarean section rate taking into account the case mix of a maternity hospital: analysis from the AUDIPOG Sentinelle Network (France). *BJOG* 2001; 108(9): 919-926.
17. Davis D. Does CME work? An analysis of the effect of educational activities on physician performance or health care outcomes. *Int J Psychiatry Med* 1998; 28(1): 21-39.
18. DeMott RK, Sandmire HF. The Green Bay cesarean section study; I: the physician factor as a determinant of cesarean birth rates. *Am J Obstet Gynecol* 1990; 162(6): 1593-1599.
19. DeMott RK, Sandmire HF. The Green Bay cesarean section study; II: the physician factor as a determinant of cesarean birth rates for failed labor. *Am J Obstet Gynecol* 1992; 166(6 Pt 1): 1799-1806.
20. Escuriet R, Pueyo M, Biescas H, Colls C, Espiga I, White J et al. Obstetric interventions in two groups of hospitals in Catalonia: a cross-sectional study. *BMC Pregnancy Childbirth* 2014; 14: 143.
21. Escuriet R, Pueyo MJ, Perez-Botella M, Espada X, Salgado I, Gomez A et al. Cross-sectional study comparing public and private hospitals in Catalonia: is the practice of routine episiotomy changing? *BMC Health Serv Res* 2015; 15: 95.
22. Friedman AM, Ananth CV, Huang Y, D'Alton ME, Wright JD. Hospital delivery volume, severe obstetrical morbidity, and failure to rescue. *Am J Obstet Gynecol* 2016; 215(6): 795.e1-795.e14.

23. Goldman G, Pineault R, Bilodeau H, Blais R. Effects of patient, physician and hospital characteristics on the likelihood of vaginal birth after previous cesarean section in Quebec. *CMAJ* 1990; 143(10): 1017-1024.
24. Goldman G, Pineault R, Potvin L, Blais R, Bilodeau H. Factors influencing the practice of vaginal birth after cesarean section. *Am J Public Health* 1993; 83(8): 1104-1108.
25. Goode KT, Weiss PM, Koller C, Kimmel S, Hess LW. Episiotomy rates in private vs. resident service deliveries: a comparison. *J Reprod Med* 2006; 51(3): 190-192.
26. Gossett DR, Dunsmoor Su R. Episiotomy practice in a community hospital setting. *J Reprod Med* 2008; 53(10): 803-808.
27. Griffiths AN, Avasarala S, Wiener JJ. A prospective observational study of emergency caesarean section rates and the effect of the labour ward experience. *J Obstet Gynaecol* 2005; 25(7): 666-668.
28. Grobman WA, Feinglass J, Murthy S. Are the Agency for Healthcare Research and Quality obstetric trauma indicators valid measures of hospital safety? *Am J Obstet Gynecol* 2006; 195(3): 868-874.
29. Han KT, Kim SJ, Ju YJ, Choi JW, Park EC. Do hospital characteristics influence cesarean delivery? Analysis of national health insurance claim data. *Eur J Public Health* 2017; 27(5): 801-807.
30. Heller G, Richardson DK, Schnell R, Misselwitz B, Künzel W, Schmidt S. Are we regionalized enough? Early-neonatal deaths in low-risk births by the size of delivery units in Hesse, Germany 1990-1999. *Int J Epidemiol* 2002; 31(5): 1061-1068.
31. Heller G, Schnell R, Richardson DK, Misselwitz B, Schmidt S. Hat die Größe der Geburtsklinik Einfluss auf das neonatale Überleben? Schätzung von „vermeidbaren“ Todesfällen in Hessen 1990–2000. *Dtsch Med Wochenschr* 2003; 128(13): 657-662.
32. Hueston WJ. Factors associated with the use of episiotomy during vaginal delivery. *Obstet Gynecol* 1996; 87(6): 1001-1005.
33. Janakiraman V, Lazar J, Joynt KE, Jha AK. Hospital volume, provider volume, and complications after childbirth in U.S. hospitals. *Obstet Gynecol* 2011; 118(3): 521-527.
34. Joyce R, Webb R, Peacock J. Predictors of obstetric intervention rates: case-mix, staffing levels and organisational factors of hospital of birth. *J Obstet Gynaecol* 2002; 22(6): 618-625.
35. Joyce R, Webb R, Peacock JL. Associations between perinatal interventions and hospital stillbirth rates and neonatal mortality. *Arch Dis Child Fetal Neonatal Ed* 2004; 89(1): F51-F56.
36. Klein MC, Spence A, Kaczorowski J, Kelly A, Grzybowski S. Does delivery volume of family physicians predict maternal and newborn outcome? *CMAJ* 2002; 166(10): 1257-1263.
37. Konstantiniuk P, Kern I, Giuliani A, Kainer F. The midwife factor in obstetric procedures and neonatal outcome. *J Perinat Med* 2002; 30(3): 242-249.

38. Kozhimannil KB, Hung P, Prasad S, Casey M, McClellan M, Moscovice IS. Birth volume and the quality of obstetric care in rural hospitals. *J Rural Health* 2014; 30(4): 335-343.
39. Kozhimannil KB, Thao V, Peiyin H, Tilden E, Caughey AB, Snowden JM. Association between hospital birth volume and maternal morbidity among low-risk pregnancies in rural, urban, and teaching hospitals in the United States. *Am J Perinatol* 2016; 33(6): 590-599.
40. Kruse L, Denk CE, Feldman-Winter L, Rotondo FM. Comparing sociodemographic and hospital influences on breastfeeding initiation. *Birth* 2005; 32(2): 81-85.
41. Kyser KL, Lu X, Santillan D, Santillan M, Caughey AB, Wilson MC et al. Forceps delivery volumes in teaching and nonteaching hospitals: are volumes sufficient for physicians to acquire and maintain competence? *Acad Med* 2014; 89(1): 71-76.
42. Labrecque M, Baillargeon L, Dallaire M, Tremblay A, Pinault JJ, Gingras S. Association between median episiotomy and severe perineal lacerations in primiparous women. *CMAJ* 1997; 156(6): 797-802.
43. Le Ray C, Carayol M, Zeitlin J, Breart G, Goffinet F. Level of perinatal care of the maternity unit and rate of cesarean in low-risk nulliparas. *Obstet Gynecol* 2006; 107(6): 1269-1277.
44. Le Ray C, Zeitlin J, Jarreau PH, Breart G, Goffinet F. The influence of level of care on admission to neonatal care for babies of low-risk nullipara. *Eur J Obstet Gynecol Reprod Biol* 2009; 144(1): 21-26.
45. LeFevre M, Sanner L, Anderson S, Tsutakawa R. The relationship between neonatal mortality and hospital level. *J Fam Pract* 1992; 35(3): 259-264.
46. Lin HC, Sheen TC, Tang CH, Kao S. Association between maternal age and the likelihood of a cesarean section: a population-based multivariate logistic regression analysis. *Acta Obstet Gynecol Scand* 2004; 83(12): 1178-1183.
47. Lin HC, Xirasagar S. Institutional factors in cesarean delivery rates: policy and research implications. *Obstet Gynecol* 2004; 103(1): 128-136.
48. Linton A, Peterson MR, Williams TV. Clinical case mix adjustment of cesarean delivery rates in U.S. military hospitals, 2002. *Obstet Gynecol* 2005; 105(3): 598-606.
49. Lundsberg LS, Illuzzi JL, Gariepy AM, Sheth SS, Pettker CM, Lee HC et al. Variation in hospital intrapartum practices and association with cesarean rate. *J Obstet Gynecol Neonatal Nurs* 2017; 46(1): 5-17.
50. Luthy DA, Malmgren JA, Zingheim RW, Leininger CJ. Physician contribution to a cesarean delivery risk model. *Am J Obstet Gynecol* 2003; 188(6): 1579-1585.
51. Maeda E, Ishihara O, Tomio J, Sato A, Terada Y, Kobayashi Y et al. Cesarean section rates and local resources for perinatal care in Japan: a nationwide ecological study using the national database of health insurance claims. *J Obstet Gynaecol Res* 2018; 44(2): 208-216.

52. Mangin M, Ramanah R, Aouar Z, Courtois L, Collin A, Cossa S et al. Operative delivery data in France for 2007: results of a national survey within teaching hospitals [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 2010; 39(2): 121-132.
53. Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. *Am J Public Health* 1990; 80(7): 819-823.
54. McClelland S, Gorfinkle N, Arslan AA, Benedetto-Anzai MT, Cheon T, Anzai Y. Factors associated with cesarean delivery rates: a single-institution experience. *Matern Health Neonatol Perinatol* 2017; 3: 8.
55. Merlo J, Gerdtham UG, Eckerlund I, Hakansson S, Otterblad-Olausson P, Pakkanen M et al. Hospital level of care and neonatal mortality in low- and high-risk deliveries: reassessing the question in Sweden by multilevel analysis. *Med Care* 2005; 43(11): 1092-1100.
56. Miller ES, Barber EL, McDonald KD, Gossett DR. Association between obstetrician forceps volume and maternal and neonatal outcomes. *Obstet Gynecol* 2014; 123(2 Pt 1): 248-254.
57. Mizrachi Y, Leytes S, Levy M, Hiaev Z, Ginath S, Bar J et al. Does midwife experience affect the rate of severe perineal tears? *Birth* 2017; 44(2): 161-166.
58. Naiditch M, Levy G, Chale JJ, Cohen H, Colladon B, Maria B et al. Cesarean sections in France: impact of organizational factors on different utilization rates [Französisch]. *J Gynecol Obstet Biol Reprod (Paris)* 1997; 26(5): 484-495.
59. Nicholson WK, Witter F, Powe NR. Effect of hospital setting and volume on clinical outcomes in women with gestational and type 2 diabetes mellitus. *J Womens Health* 2009; 18(10): 1567-1576.
60. Parazzini F, Cipriani S, Bulfoni G, Bulfoni C, Bellu R, Zanini R et al. Mode of delivery and level of neonatal care in Lombardy: a descriptive analysis according to volume of care. *Ital J Pediatr* 2015; 41: 24.
61. Phibbs CS, Bronstein JM, Buxton E, Phibbs RH. The effects of patient volume and level of care at the hospital of birth on neonatal mortality. *JAMA* 1996; 276(13): 1054-1059.
62. Poets CF, Bartels DB, Wallwiener D. Volumen- und Ausstattungsmerkmale als peri- und neonatale Qualitätsindikatoren: eine Übersicht über Daten der letzten 4 Jahre. *Z Geburtshilfe Neonatol* 2004; 208(6): 220-225.
63. Pyykönen A, Gissler M, Jakobsson M, Lehtonen L, Tapper AM. The rate of obstetric anal sphincter injuries in Finnish obstetric units as a patient safety indicator. *Eur J Obstet Gynecol Reprod Biol* 2013; 169(1): 33-38.
64. Pyykönen A, Gissler M, Jakobsson M, Petäjä J, Tapper AM. Determining obstetric patient safety indicators: the differences in neonatal outcome measures between different-sized delivery units. *BJOG* 2014; 121(4): 430-437.

65. Roman H, Carayol M, Watier L, Le Ray C, Breart G, Goffinet F. Planned vaginal delivery of fetuses in breech presentation at term: prenatal determinants predictive of elevated risk of cesarean delivery during labor. *Eur J Obstet Gynecol Reprod Biol* 2008; 138(1): 14-22.
66. Sagot P, Gouyon-Cornet B, Gouyon JB. Cesarean section in the Burgundy perinatal network: reasons for differences in the management of mothers [Französisch]. *Gynecol Obstet Fertil* 2003; 31(7-8): 606-613.
67. Sandall J, Murrells T, Dodwell M, Gibson R, Bewley S, Coxon K et al. The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study. *Health Serv Deliv Res* 2014; 2(38).
68. Sandmire HF, DeMott RK. The Green Bay cesarean section study; III: falling cesarean birth rates without a formal curtailment program. *Am J Obstet Gynecol* 1994; 170(6): 1790-1799.
69. Sandmire HF, DeMott RK. The Green Bay cesarean section study; IV: the physician factor as a determinant of cesarean birth rates for the large fetus. *Am J Obstet Gynecol* 1996; 174(5): 1557-1564.
70. Simoes E, Kunz S, Bosing-Schwenkglens M, Schmahl FW. Mütterliche Anämie post partum: Entwicklungstendenzen und Variabilität unter Berücksichtigung unterschiedlicher Klinikategorien; Untersuchung auf der Basis der Perinatalerhebung in Baden-Württemberg. *Z Geburtshilfe Neonatol* 2004; 208(5): 184-189.
71. Snowden JM, Cheng YW, Emeis CL, Caughey AB. The impact of hospital obstetric volume on maternal outcomes in term, non-low-birthweight pregnancies. *Am J Obstet Gynecol* 2015; 212(3): 380.e1-380.e9.
72. Snowden JM, Cheng YW, Kontgis CP, Caughey AB. The association between hospital obstetric volume and perinatal outcomes in California. *Am J Obstet Gynecol* 2012; 207(6): 478.e1-478.e7.
73. Stafford RS. The impact of nonclinical factors on repeat cesarean section. *JAMA* 1991; 265(1): 59-63.
74. Takacs L, Seidlerova J. Psychosocial climate in maternity hospitals from the perspective of parturients I: results from a national survey on perinatal care satisfaction using a representative sample of 1195 Czech parturients [Tschechisch]. *Ceska Gynekol* 2013; 78(2): 157-168.
75. Tracy SK, Sullivan E, Dahlen H, Black D, Wang YA, Tracy MB. Does size matter? A population-based study of birth in lower volume maternity hospitals for low risk women. *BJOG* 2006; 113(1): 86-96.
76. Vankan E, Schoorel EN, Van Kuijk SM, Mol BJ, Nijhuis JG, Aardenburg R et al. Practice variation of vaginal birth after cesarean and the influence of risk factors at patient level: a retrospective cohort study. *Acta Obstet Gynecol Scand* 2017; 96(2): 158-165.

77. Vendittelli F, Barasinski C, Pereira B, Lemery D. Incidence of immediate postpartum hemorrhages in French maternity units: a prospective observational study (HERA study). *BMC Pregnancy Childbirth* 2016; 16: 242.

78. Viisainen K, Gissler M, Hemminki E. Birth outcomes by level of obstetric care in Finland: a catchment area based analysis. *J Epidemiol Community Health* 1994; 48(4): 400-405.

Nicht E6 (Publikation)

1. Bond S. Study suggests providers attending high volumes of births have lower rates of complications, but flaws and limitations render findings ambiguous. *J Midwifery Womens Health* 2012; 57(1): 97-98.

2. Lenguerrand E, Winter C, Innes K, MacLennan G, Siassakos D, Lynch P et al. THISTLE: trial of hands-on Interprofessional simulation training for local emergencies: a research protocol for a stepped-wedge clustered randomised controlled trial. *BMC Pregnancy Childbirth* 2017; 17(1): 294.

3. Romijn A, De Bruijne MC, Teunissen PW, De Groot CJM, Wagner C. Complex social intervention for multidisciplinary teams to improve patient referrals in obstetrical care: protocol for a stepped wedge study design. *BMJ Open* 2016; 6(7): e011443.

4. Van de Ven J, Houterman S, Steinweg RAJQ, Scherpbier AJJA, Wijers W, Mol BWJ et al. Reducing errors in health care: cost-effectiveness of multidisciplinary team training in obstetric emergencies (TOSTI study); a randomised controlled trial. *BMC Pregnancy Childbirth* 2010; 10: 59.

5. Vintzileos AM. Measuring the impact of attending physician teaching in an obstetrics and gynecology residency program. *Am J Obstet Gynecol* 2017; 217(5): 617-618.

6. Walker D, Cohen S, Fritz J, Olvera M, Zelek S, Greenberg Cowan J et al. PRONTO: obstetric and neonatal emergency simulation in Mexico improves patient outcomes, provider knowledge, team coordination, and identifies latent systems errors. *J Midwifery Womens Health* 2014; 59(5): 548-549.

Anhang A – Suchstrategien**A.1 – Bibliografische Literaturrecherche****1. MEDLINE: Übergeordnete Suchstrategie zur Erfahrung / Expertise*****Übergeordnete Suchstrategie zur Erfahrung / Expertise****Suchoberfläche: Ovid*

- Ovid MEDLINE(R) 1946 to February Week 4 2018
- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations March 06, 2018
- Ovid MEDLINE(R) Daily Update March 06, 2018
- Ovid MEDLINE(R) Epub Ahead of Print March 06, 2018

#	Searches
1	*Delivery, Obstetric/
2	*Obstetrics/
3	(deliveries* or birth?).ti,ab.
4	or/1-3
5	*Perinatal Care/
6	exp Neonatology/
7	"Hospitals, Maternity"/
8	*"Obstetrics and Gynecology Department, Hospital"/
9	or/5-8
10	9 and level?.ti,ab.
11	(level? adj1 (I or II or III) adj3 hospital?).ti,ab.
12	(level? adj3 ((neonatal* or perinatal*) adj3 (care or service?))).ti,ab.
13	or/10-12
14	Clinical Competence/
15	14 and (midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist?).ti,ab.
16	((characteristic* or factor?) adj3 (midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist? or organi#ational?)).ti,ab.
17	or/15-16
18	exp *"Health Facility Size"/
19	9 and (size? or characteristic*).ti,ab.
20	"Practice Patterns, Physicians"/sn
21	or/18-20
22	or/13,17,21
23	and/4,22

#	Searches
24	((small* or large? or size? or characteristic*) adj3 (maternity* or delivery* or obstetric*) adj3 unit?).ti,ab.
25	((hospital? or unit? or midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist? or provider?) adj5 ((deliveries* or delivery* or births*) adj3 (year? or number? or annu* or volume? or case?))).ti,ab.
26	or/24-25
27	or/23,26
28	27 not (exp animals/ not humans.sh.)
29	28 not (comment or editorial).pt.
30	..l/ 29 yr=1990-Current

Spezifische Suchstrategie zu Weiterbildungsmaßnahmen

Suchoberfläche: Ovid

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations September 28, 2018,
- Ovid MEDLINE(R) 1946 to September Week 3 2018,
- Ovid MEDLINE(R) Daily Update September 28, 2018,
- Ovid MEDLINE(R) Epub Ahead of Print September 28, 2018

Es wurden folgende Filter übernommen:

- Systematische Übersicht: Wong [116] – High specificity strategy
- RCT: Lefebvre [117] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision)

#	Searches
4	Pregnancy/
5	Delivery, Obstetric/
6	(labor* or deliver* or birth* or childbirth*).ab,ti,kw.
7	or/4-6
8	(midwi* or nurse? or physician? or obstetric* or consultant? or neonatologist?).ti,ab,kw.
9	Cesarean Section/
10	8 or 9
11	Medical Audit/
12	Quality of Health Care/
13	Clinical Competence/
14	Peer Review, Health Care/

#	Searches
15	Guideline Adherence/
16	Education, Medical, Continuing/
17	or/11-16
18	(ed or sn or ut).fs.
19	and/17-18
20	(((clinical* or medical*) adj1 audit*) or ((detail* or provision* or written*) adj3 feedback*) or opinion leader*).ab,ti,kw.
21	(second adj3 opinion).ab,ti,kw.
22	(simulation* adj5 training).ab,ti,kw.
23	or/19-22
24	randomized controlled trial.pt.
25	controlled clinical trial.pt.
26	randomized.ab.
27	placebo.ab.
28	clinical trial as topic/
29	randomly.ab.
30	trial.ti.
31	or/24-30
32	exp animals/ not humans.sh.
33	31 not 32
34	Cochrane database of systematic reviews.jn.
35	(search or MEDLINE or systematic review).tw.
36	meta analysis.pt.
37	or/34-36
38	or/33,37
39	and/7,10,23,38
40	39 not (comment or editorial).pt.
41	../ 40 yr=1990-Current

2. Embase

Suchoberfläche: Ovid

- Embase 1974 to 2018 March 06

Es wurden folgende Filter übernommen:

- Systematische Übersicht: Wong [116] – High specificity strategy;

- RCT: Wong [116] – Strategy minimizing difference between sensitivity and specificity

#	Searches
1	delivery/
2	exp obstetrics/
3	birth/
4	*labor/
5	obstetric patient/
6	(deliveries* or birth?).ti,ab.
7	or/1-6
8	perinatal care/
9	newborn intensive care/
10	maternity ward/
11	delivery room/
12	or/8-11
13	12 and level?.ti,ab.
14	(level? adj1 (I or II or III) adj3 hospital?).ti,ab.
15	(level? adj3 ((neonatal* or perinatal*) adj3 (care or service?))).ti,ab.
16	or/13-15
17	exp medical education/
18	17 and (midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist?).ti,ab.
19	((characteristic* or factor?) adj3 (midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist? or organi#ational?)).ti,ab.
20	or/18-19
21	12 and (size? or characteristic*).ti,ab.
22	or/16,20-21
23	and/7,22
24	((small* or large? or size? or characteristic*) adj3 (maternity* or delivery* or obstetric*) adj3 unit?).ti,ab.
25	((hospital? or unit? or midwi* or nurse? or physician? or obstetrician? or consultant? or neonatologist? or provider?) adj5 ((deliveries* or delivery* or births*) adj3 (year? or number? or annu* or volume? or case?))).ti,ab.
26	or/24-25
27	or/23,26
28	(random* or double-blind*).tw.
29	placebo*.mp.
30	or/28-29

#	Searches
31	(meta analysis or systematic review or MEDLINE).tw.
32	or/30-31
33	and/27,32
34	33 not medline.cr.
35	34 not (exp animal/ not exp humans/)
36	35 not (Conference Abstract or Conference Review or Editorial).pt.
37	..l/ 36 yr=1990-Current

3. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews: Issue 3 of 12, March 2018
- Cochrane Central Register of Controlled Trials: Issue 2 of 12, February 2018

ID	Search
#1	[mh ^"Delivery, Obstetric"]
#2	[mh ^Obstetrics]
#3	(deliveries* or birth*):ti,ab
#4	#1 or #2 or #3
#5	[mh ^"Perinatal Care"]
#6	[mh Neonatology]
#7	[mh ^"Hospitals, Maternity"]
#8	[mh ^"Obstetrics and Gynecology Department, Hospital"]
#9	#5 or #6 or #7 or #8
#10	#9 and level*:ti,ab
#11	(level* near/1 (I or II or III) near/3 hospital*):ti,ab
#12	(level* near/3 ((neonatal* or perinatal*) near/3 (care or service*))) :ti,ab
#13	#10 or #11 or #12
#14	[mh ^"Clinical Competence"]
#15	#14 and (midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist*):ti,ab
#16	((characteristic* or factor*) near/3 (midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist* or organi?ational*)):ti,ab
#17	#15 or #16
#18	[mh "Health Facility Size"]
#19	#9 and (size* or characteristic*):ti,ab

ID	Search
#20	MeSH descriptor: [Practice Patterns, Physicians'] this term only and with qualifier(s): [Statistics & numerical data - SN]
#21	#18 or #19 or #20
#22	#13 or #17 or #21
#23	#4 and #22
#24	((small* or large* or size* or characteristic*) near/3 (maternity* or delivery* or obstetric*) near/3 unit*):ti,ab
#25	((hospital* or unit* or midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist* or provider*) near/5 ((deliveries* or delivery* or births*) near/3 (year* or number* or annu* or volume* or case*)):ti,ab
#26	#24 or #25
#27	#23 or #26
#28	#27 in Cochrane Reviews (Reviews and Protocols)
#29	#27 Publication Year from 1990 to 2018, in Trials

4. Health Technology Assessment Database

Suchoberfläche: Centre for Reviews and Dissemination

Line	Search
1	MeSH DESCRIPTOR Delivery, Obstetric
2	MeSH DESCRIPTOR Obstetrics
3	(deliveries* or birth*)
4	#1 OR #2 OR #3
5	MeSH DESCRIPTOR Perinatal Care
6	MeSH DESCRIPTOR Neonatology EXPLODE ALL TREES
7	MeSH DESCRIPTOR Hospitals, Maternity
8	MeSH DESCRIPTOR Obstetrics and Gynecology Department, Hospital
9	#5 OR #6 OR #7 OR #8
10	(#9 and level*)
11	(level* and (I or II or III) and hospital*)
12	(level* and (neonatal* or perinatal*) and (care or service*))
13	#10 OR #11 OR #12
14	MeSH DESCRIPTOR Clinical Competence
15	(#14 and (midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist*))
16	((characteristic* or factor*) AND (midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist* or organizational* or organisational*))

Line	Search
17	#15 OR #16
18	MeSH DESCRIPTOR Health Facility Size EXPLODE ALL TREES
19	(#9 and (size* or characteristic*))
20	MeSH DESCRIPTOR Practice Patterns, Physicians' EXPLODE ALL TREES
21	#18 OR #19 OR #20
22	#13 OR #17 OR #21
23	#4 AND #22
24	((small* or large* or size* or characteristic*) and (maternity* or delivery* or obstetric*) and unit*)
25	((hospital* or unit* or midwi* or nurse* or physician* or obstetrician* or consultant* or neonatologist* or provider*) and (deliveries* or delivery* or births*) and (year* or number* or annu* or volume* or case*))
26	#24 OR #25
27	#23 OR #26
28	(#27) IN HTA FROM 1990 TO 2018

5. Cinahl

Suchoberfläche: Ebsco

#	Query
S1	(MH "Delivery, Obstetric")
S2	(MM "Obstetrics")
S3	TI (deliveries* or birth#) OR AB (deliveries* or birth#)
S4	S1 OR S2 OR S3
S5	(MH "Intensive Care, Neonatal")
S6	(MM "Intensive Care Units, Neonatal")
S7	(MM "Perinatal Care")
S8	S5 OR S6 OR S7
S9	S8 AND (TI level# OR AB level#)
S10	TI (level# N3 (I OR II OR III) N3 hospital#) OR AB (level# N3 (I OR II OR III) N3 hospital#)
S11	TI (level# N3 ((neonatal* OR perinatal*) N3 (care OR service#))) OR AB (level# N3 ((neonatal* OR perinatal*) N3 (care OR service#)))
S12	S9 OR S10 OR S11
S13	(MH "Clinical Competence+")
S14	(MW "ED")
S15	S13 OR S14

#	Query
S16	S15 AND (TI (midwi* OR nurse# OR physician# OR obstetrician# OR consultant# OR neonatologist#) OR AB (midwi* OR nurse# OR physician# OR obstetrician# OR consultant# OR neonatologist#))
S17	TI ((characteristic* OR factor#) N3 (midwi* OR nurse# OR physician# OR obstetrician# OR Consultant# OR neonatologist# OR organisational* OR organizational*)) OR AB ((characteristic* OR factor#) N3 (midwi* OR nurse# OR physician# OR obstetrician# OR consultant# OR neonatologist# OR organisational* OR organizational*))
S18	S16 OR S17
S19	S8 AND (TI (size# or characteristic*) OR AB (size# or characteristic*))
S20	S12 OR S18 OR S19
S21	S4 AND S20
S22	TI ((small* OR large# OR size# OR characteristic*) N3 (maternity* OR delivery* OR obstetric*) N3 unit#) OR AB ((small* OR large# OR size# OR characteristic*) N3 (maternity* OR delivery* OR obstetric*) N3 unit#)
S23	TI ((hospital# OR unit# OR midwi* OR nurse# OR physician# OR obstetrician# OR consultant# OR neonatologist# OR provider#) N5 ((deliveries* OR delivery* OR births*) N3 (year# OR number# OR annu* OR volume# OR case#))) OR AB ((hospital# OR unit# OR midwi* OR nurse# OR physician# OR obstetrician# OR consultant# OR neonatologist# OR provider#) N5 ((deliveries* OR delivery* OR births*) N3 (year# OR number# OR annu* OR volume# OR case#)))
S24	S22 OR S23
S25	S21 OR S24
S26	S25 AND (PY 1990-2018)

A.2 – Suche in Studienregistern

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: <http://www.clinicaltrials.gov>
- Eingabeoberfläche: Advanced Search

Suchstrategie
birth AND ("level I" OR "level II" OR "level III" OR "neonatal service" OR "neonatal center" OR "perinatal service" OR "perinatal center")
(characteristic OR factor) AND (midwife OR midwifery OR nurse OR physician OR obstetrician OR consultant OR neonatologist OR organizational) AND birth [DISEASE]

2. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <http://apps.who.int/trialsearch/>
- Eingabeoberfläche: Standard Search

Suchstrategie
level AND birth AND hospital OR level AND birth* AND hospital OR level AND delivery* AND hospital OR neonatal cent* OR neonatal service* OR perinatal service* OR perinatal cent*
characteristic* AND birth OR characteristic* AND birth* OR characteristic* AND delivery* OR factor* AND birth OR factor* AND birth* OR factor* AND delivery*

Anhang B – Verzerrungspotenzial auf Endpunktebene

B.1 – Erfahrung / Expertise durch Weiterbildungsmaßnahmen auf der Grundlage von randomisierten kontrollierten Studien

B.1.1 – Maternale Mortalität

Die folgende Tabelle 20 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für die maternale Mortalität.

Tabelle 20: Bewertung des Verzerrungspotenzials auf Endpunktebene zu maternale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	ja	ja	ja	nein	gering ^a
Lomas 1991	hoch	ja	ja	ja	nein	hoch

a: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt).
ITT: Intention to treat

B.1.2 – Geburtsmodus

Die folgende Tabelle 21 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für den Geburtsmodus.

Tabelle 21: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^{a, b}
Lomas 1991	hoch	unklar	unklar	ja	nein	hoch ^{a, b, c}

a: Fehlende oder unklare Verblindung der behandelnden oder endpunkterhebenden Personen kann einen Einfluss auf die Erfassung und Differenzierung eines geplanten oder ungeplanten Kaiserschnitts haben.
b: Effektschätzer wurden für einige Ausprägungen ohne Berücksichtigung der Clusterstruktur berechnet.
c: intransparenter Patientenfluss: Angaben absoluter Zahlen fehlen; widersprüchliche Angaben / Dateninkonsistenzen bezüglich der Prozentangaben zu Frauen mit einem Geburtsversuch, vaginaler Geburt und ungeplantem Kaiserschnitt im Geburtsverlauf
ITT: Intention to treat

B.1.3 – Höhergradige Dammrisse (III°, IV°)

Die folgende Tabelle 22 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für höhergradige Dammrisse (III°, IV°).

Tabelle 22: Bewertung des Verzerrungspotenzials auf Endpunktebene zu höhergradige Dammrisse (III°, IV°) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^{a, b}

a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen
b: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt).
ITT: Intention to treat

B.1.4 – Neonatale Mortalität

Die folgende Tabelle 23 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für die neonatale Mortalität.

Tabelle 23: Bewertung des Verzerrungspotenzials auf Endpunktebene zu neonataler Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	ja	ja	ja	nein	mittel ^{a, b}
Lomas 1991	hoch	ja	ja	ja	ja	hoch

a: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt).
b: Effektschätzer wurde ohne Berücksichtigung der Clusterstruktur berechnet
ITT: Intention to treat

B.1.5 – Schwere metabolische Acidose

Die folgende Tabelle 24 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für die schwere metabolische Acidose.

Tabelle 24: Bewertung des Verzerrungspotenzials auf Endpunktebene zu schwerer metabolischer Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	gering ^{a, b}

a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen bei objektiv erhobenem Endpunkt nicht relevant
b: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt).
ITT: Intention to treat

B.1.6 – Episiotomie

Die folgende Tabelle 25 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für Episiotomie.

Tabelle 25: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Episiotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^a
a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen ITT: Intention to treat						

B.1.7 – Wehenmittel

Die folgenden Tabellen Tabelle 26 und Tabelle 27 zeigen die endpunktspezifische Bewertung des Verzerrungspotenzials für Wehenmittelgabe unter der Geburt und zur Einleitung der Geburt.

Tabelle 26: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Wehenmitteln unter der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^a
a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen ITT: Intention to treat						

Tabelle 27: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Wehenmittelgabe zur Einleitung der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^a
a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen ITT: Intention to treat						

B.1.8 – Amniotomie

Die folgende Tabelle 28 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für Amniotomie.

Tabelle 28: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Amniotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	mittel ^{a, b}
a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen b: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt). ITT: Intention to treat						

B.1.9 – Infektionen

Die folgende Tabelle 29 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für Infektionen.

Tabelle 29: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Infektion zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillet 2015	gering	nein	ja	ja	nein	gering ^a
a: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt). ITT: Intention to treat						

B.1.10 – Leichte, mittelgradige und fortgeschrittene Acidose

Die folgende Tabelle 30 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für fortgeschrittene Acidose.

Tabelle 30: Bewertung des Verzerrungspotenzials auf Endpunktebene zu fortgeschrittener Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chaillot 2015	gering	nein	ja	ja	nein	gering ^{a, b, c}
a: fehlende Verblindung der behandelnden oder endpunkterhebenden Personen bei objektiv erhobenem Endpunkt nicht relevant b: Ergebnis basiert auf allen Gebärenden (einschließlich der Frauen mit geplantem Kaiserschnitt). c: Effektschätzer wurde ohne Berücksichtigung der Clusterstruktur berechnet ITT: Intention to treat						

B.2 – Erfahrung / Expertise durch systemimmanente Faktoren

B.2.1 – Geburtsmodus

Die folgende Tabelle 31 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für den Geburtsmodus.

Tabelle 31: Bewertung des Verzerrungspotenzials auf Endpunktebene zu Geburtsmodus zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Coulm 2012	hoch	ja	ja	ja	ja	hoch ^a
Wehberg 2018	hoch	ja	ja	ja	ja	hoch ^a
a: Bewertung basiert auf der Bewertung auf Studienebene. ITT: Intention to treat						

B.2.2 – Neonatale Mortalität

Die folgende Tabelle 32 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für die neonatale Mortalität.

Tabelle 32: Bewertung des Verzerrungspotenzials auf Endpunktebene zu neonataler Mortalität zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Finnström 2006 ^a	hoch	ja	unklar	ja	ja	hoch ^b
Karalis 2016 ^c	hoch	ja	unklar	ja	ja	hoch ^{b, d}

a: Berichtet werden neonatale Todesfälle (1.–27. Lebenstag).
b: Bewertung basiert auf der Bewertung auf Studienebene.
c: Berichtet werden intrapartale sowie frühe neonatale Todesfälle (1.–7. Lebenstag).
d: Daten zu 324 (0,12 %) der Frauen bzw. Neugeborenen wurden aus der Analyse ausgeschlossen. Die fehlenden Werte können die Ergebnisse zu diesem seltenen Ereignis beeinflussen.
ITT: Intention to treat

B.2.3 – Schwere metabolische Acidose (pH < 7,0)

Die folgende Tabelle 33 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für schwere metabolische Acidose.

Tabelle 33: Bewertung des Verzerrungspotenzials auf Endpunktebene: schwere metabolische Acidose zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Karalis 2016	hoch	ja	unklar	ja	ja	hoch ^{a, b}

a: Daten zu 324 (0,12 %) der Frauen bzw. Neugeborenen wurden aus der Analyse ausgeschlossen. Die fehlenden Werte können die Ergebnisse zu diesem seltenen Ereignis beeinflussen.
b: Bewertung basiert auf der Bewertung auf Studienebene.
ITT: Intention to treat

B.2.4 – Episiotomie

Die folgende Tabelle 34 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für die Episiotomie.

Tabelle 34: Bewertung des Verzerrungspotenzials auf Endpunktebene: Episiotomie zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Chuilon 2016	hoch	unklar	ja	ja	ja	hoch ^{a, b}
Coulm 2012	hoch	unklar	ja	ja	ja	hoch ^{a, b}

a: Es liegen keine Informationen zur Erfassung und Datenextraktion aus den Krankenhausakten dieses subjektiv beeinflussbaren Endpunktes vor.
b: Bewertung basiert auf der Bewertung auf Studienebene.
ITT: Intention to treat

B.2.5 – Wehenmittel

Die folgende Tabelle 35 zeigt die endpunktspezifische Bewertung des Verzerrungspotenzials für Wehenmittelgabe.

Tabelle 35: Bewertung des Verzerrungspotenzials auf Endpunktebene: Wehenmittel zur Erfahrung / Expertise durch systemimmanente Faktoren

Studie	Verzerrungspotenzial auf Studienebene	Verblindung endpunkterhebende Personen	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Endpunktebene
Belghiti 2013 ^a	hoch	unklar	ja	ja	ja	hoch ^{b, c}

a: Berichtet wird die Gabe von Oxytocin während der Geburt bei Gebärenden mit französischer Standardbetreuung.
b: Es liegen keine Informationen zur Erfassung und Datenextraktion aus den Krankenhausakten dieses subjektiv beeinflussbaren Endpunktes vor.
c: Bewertung basiert auf der Bewertung auf Studienebene.
ITT: Intention to treat

Anhang C – Ergebnisse

C.1 – Ergebnisse zu Erfahrung / Expertise durch Weiterbildungsmaßnahmen

In diesem Abschnitt folgt eine tabellarische Darstellung der Ergebnisse zu den kritischen und wichtigen Endpunkten für die Suche nach vergleichenden Interventionsstudien zur Erfahrung / Expertise des geburtshilflichen Fachpersonals. Aus der Kategorie der kritischen Endpunkte wurden Ergebnisse zur maternalen Mortalität (siehe Tabelle 36), Geburtsmodus (siehe Tabelle 37 und Tabelle 38), zur neonatalen Mortalität (siehe Tabelle 40), zu höhergradigen Dammrissen (siehe Tabelle 39) und schwerer metabolischer Acidose (siehe Tabelle 43) berichtet. Zusätzlich werden ergänzende Informationen zur neonatalen Mortalität (siehe Tabelle 41) und zum Apgar-Score nach 5 Minuten (siehe Tabelle 42) berichtet.

Aus der Kategorie der wichtigen Endpunkte wurden Ergebnisse zur Episiotomie (siehe Tabelle 44), Wehenmittel (siehe Tabelle 45 und Tabelle 46), Amniotomie (siehe Tabelle 47), Infektionen (siehe Tabelle 48 und Tabelle 49) und fortgeschrittener Acidose (siehe Tabelle 50) berichtet.

C.1.1 – Maternale Mortalität

Die folgende Tabelle 36 zeigt die Ergebnisse für die maternale Mortalität.

Tabelle 36: Ergebnisse für die maternale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Maternale Mortalität	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 47 630 ^c	3 (0,0063) ^d	2,08 [0,05; 81,6] ^e
Vergleichsintervention	16 / 57 129 ^c	1 (0,0018) ^d	RD: 0,00 [-0,00; 0,28] ^f
Lomas 1991 (Schulung von Meinungsführern)			
Prüfintervention	4 / 345 ^g	0 (0)	RD: 0,00 [-0,01; 0,01]
Vergleichsintervention	8 / 410 ^f	0 (0)	p > 0,999 ^d
Lomas 1991 (Audit und Feedback)			
Prüfintervention	4 / 159 ^g	0 (0)	RD: 0,00 [-0,01; 0,02]
Vergleichsintervention	8 / 410 ^g	0 (0)	p > 0,999 ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Gebärender c: Anzahl in die Analyse eingeschlossener Gebärender aus der Pre- und Postinterventionsperiode d: eigene Berechnung ohne Berücksichtigung der Clusterstruktur e: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) f: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR g: Anzahl Gebärender mit vorherigem Kaiserschnitt, bei denen der Versuch einer Spontangeburt laut Leitlinie möglich war und angestrebt wurde KI: Konfidenzintervall; OR: adjustiertes Odds Ratio; n: Anzahl; RD: Risikodifferenz, SSW: Schwangerschaftswoche			

C.1.2 – Geburtsmodus

Die folgende Tabelle 37 zeigt die Ergebnisse für den Geburtsmodus.

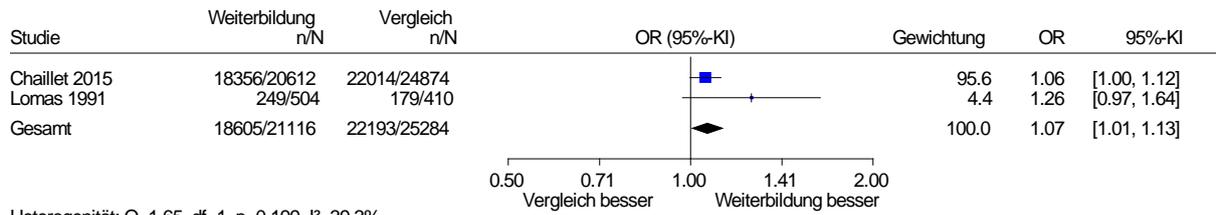
Tabelle 37: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (3 Ausprägungen des Endpunktes berichtet)

Studie (Weiterbildung)	n ^a / n ^b	Spontangeburt		Kaiserschnitt		Vaginal-operative Geburt	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI]	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI]	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)							
Prüfintervention	16 / 20 612	16 133 ^c (78,3) ^c	1,01 [0,97; 1,06] ^c / [0,96; 1,07]	2256 ^c (10,9)	0,91 [0,80; 1,03] ^d / [0,78; 1,06]	2223 (10,8)	0,88 [0,77; 0,99] ^d / [0,75; 1,03]
Vergleichsintervention	16 / 24 874	19 409 ^c (78,0) ^c		2860 ^c (11,5)	RD: -0,9 [-2,1; 0,3] ^e	2605 (10,5)	RD: -1,1 [-2,2; -0,1] ^e
<p>a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase, bei denen kein geplanter Kaiserschnitt vorgesehen war c: eigene Berechnung, Berechnung des OR ohne Berücksichtigung der Clusterstruktur d: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) e: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustiertes relatives Risiko; SSW: Schwangerschaftswoche</p>							

Tabelle 38: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen (2 Ausprägungen des Endpunktes berichtet)

Studie (Weiterbildung)	n ^a / n ^b	Vaginale Geburt ^c		Kaiserschnitt ^d	
		Gebärende mit Ereignis n (%)	RR [95 %-KI] / RR [98,75 %-KI]	Gebärende mit Ereignis n (%)	RR [95 %-KI] / RR [98,75 %-KI]
Lomas 1991 (Schulung von Meinungsführern)					
Prüfintervention	4 / 345	187 ^e (54,2)	1,24 [1,07; 1,44] ^e /	158 ^e (45,8)	0,81 [0,70; 0,94] ^e /
Vergleichsintervention	8 / 410	179 ^e (43,7)	1,24 [1,03; 1,50] ^e	231 ^e (56,3)	0,81 [0,67; 0,98] ^e
Lomas 1991 (Audit und Feedback)					
Prüfintervention	4 / 159	62 ^e (39,0)	0,89 [0,71; 1,12] ^e /	97 ^e (61,0)	1,08 [0,93; 1,26] ^e /
Vergleichsintervention	8 / 410	179 ^e (43,7)	0,89 [0,67; 1,19] ^e	231 ^e (56,3)	1,08 [0,89; 1,31] ^e
<p>a: Anzahl randomisierter Krankenhäuser b: Anzahl Gebärender mit vorherigem Kaiserschnitt, bei denen der Versuch einer Spontangeburt laut Leitlinie möglich war und angestrebt wurde c: Angegeben wird die Anzahl der vaginalen Geburten ohne Differenzierungsmöglichkeit zwischen Spontangeburt und vaginal-operativer Geburt. d: Berichtet wird die Anzahl ungeplanter und geplanter Kaiserschnitte, für diesen Bericht wurde die Anzahl ungeplanter Kaiserschnitte extrahiert. e: Eigene Berechnung. Eine Berücksichtigung der Clusterstruktur ist auf Basis der verfügbaren Daten nicht möglich. KI: Konfidenzintervall; n: Anzahl; RR: relatives Risiko</p>					

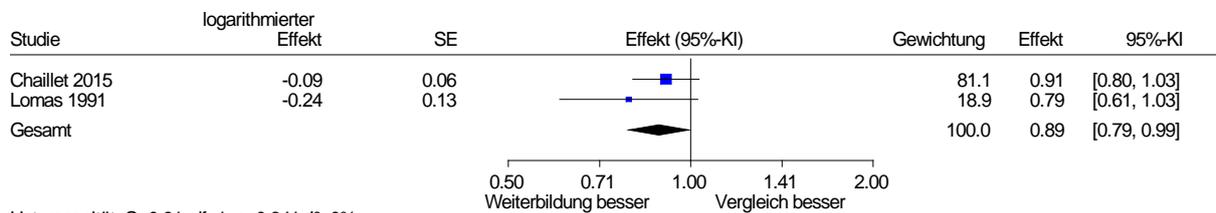
Weiterbildung vs. Vergleich
 Vaginale Geburt
 Modell mit festem Effekt - Mantel-Haenszel



Adjustiertes 97,5 % KI: [1,00; 1,14]

Abbildung 2: Forest Plot für die vaginale Geburt

Weiterbildung vs. Vergleich
 Kaiserschnitt
 Modell mit festem Effekt - Inverse Varianz



Adjustiertes 97,5 % KI: [0,78; 1,01], Effekt entspricht dem OR

Abbildung 3: Forest Plot für den Kaiserschnitt

C.1.3 – Höhergradige Dammrise (III°, IV°)

Die folgende Tabelle 39 zeigt die Ergebnisse für höhergradige Dammrise (III°, IV°).

Tabelle 39: Ergebnisse für höhergradige Dammrise (III°, IV°) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Dammriss (III°, IV°)	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 20 612	1241 (6,0)	1,07 [0,90; 1,26] ^c
Vergleichsintervention	16 / 24 874	1594 (6,4) ^d	RD: 0,4 [-0,5; 1,3] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase ohne geplanten Kaiserschnitt c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; OR: adjustiertes Odds Ratio; n: Anzahl; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.4 – Neonatale Mortalität

Tabelle 40 und Tabelle 41 zeigen die Ergebnisse für die neonatale Mortalität.

Tabelle 40: Ergebnisse für die neonatale Mortalität zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Neonatale Mortalität	
		Neugeborene mit Ereignis n (%)	RD [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm) ^c			
Prüfintervention	16 / 23 902 ^d	20 (0,08) ^e	-0,06 [-0,08; -0,03] ^f
Vergleichsintervention	16 / 29 211 ^d	28 (0,1) ^e	OR: 0,34 [0,18; 0,65] ^g p < 0,001
Lomas 1991 (Schulung von Meinungsführern)			
Prüfintervention	4 / 345 ^{e, h}	0 (0)	0,00 [-0,01; 0,01]
Vergleichsintervention	8 / 410 ^{e, h}	0 (0)	p > 0,999 ^c
Lomas 1991 (Audit und Feedback)			
Prüfintervention	4 / 159 ^{e, h}	0 (0)	0,00 [-0,01; 0,02]
Vergleichsintervention	8 / 410 ^{e, h}	0 (0)	p > 0,999 ^c
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Neugeborener c: Berichtet werden intrapartale und neonatale Todesfälle. d: Anzahl Neugeborener in der Postinterventionsphase e: eigene Berechnung ohne Berücksichtigung der Clusterstruktur f: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungssatz identisch zu OR g: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage), da das GEE-Modell nicht konvergierte konnte Clusterstruktur nicht berücksichtigt werden, ein statistischer Test erfolgte zum Signifikanzniveau von 0,001. h: Anzahl Neugeborener von Gebärenden mit vorherigem Kaiserschnitt, bei denen der Versuch einer Spontangeburt laut Leitlinie möglich war und angestrebt wurde. KI: Konfidenzintervall; OR: adjustiertes Odds Ratio; n: Anzahl; RD: Risikodifferenz; SSW: Schwangerschaftswoche			

Tabelle 41: Ergänzende Informationen für die neonatale Mortalität (Totgeburten) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Totgeburten ^c	
		Neugeborene mit Ereignis n (%)	RR [95 %-KI]
Lomas 1991 (Schulung von Meinungsführern)			
Prüfintervention	4 / 739	4 ^d (0,5) ^d	1,67 [0,42; 6,65] ^d
Vergleichsintervention	8 / 1233	4 ^d (0,3) ^d	
Lomas 1991 (Audit und Feedback)			
Prüfintervention	4 / 524	2 ^d (0,4)	1,18 [0,22; 6,40] ^d
Vergleichsintervention	8 / 1233	4 ^d (0,3)	
a: Anzahl randomisierter Krankenhäuser b: Anzahl Teilnehmerinnen mit vorherigem Kaiserschnitt, bei denen der Versuch einer Spontangeburt laut Leitlinie möglich war c: Es liegen keine Angaben zum Todeszeitpunkt vor. Laut Autorenangabe Ereignis unabhängig vom Geburtsmodus. d: eigene Berechnung ohne Berücksichtigung der Clusterstruktur KI: Konfidenzintervall; n: Anzahl; RR: relatives Risiko			

C.1.5 – Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6

Die folgende Tabelle 42 zeigt die ergänzenden Informationen zum Auftreten eines geringen Apgar-Scores nach 5 Minuten. Es lagen ausschließlich Informationen zum Apgar-Score nach 5 Minuten < 7 vor.

Tabelle 42: Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6 (Apgar-Score nach 5 Minuten < 7) zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Apgar-Score nach 5 Minuten < 7	
		Neugeborene mit Ereignis n (%)	RR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm) ^c			
Prüfintervention	16 / 23 902 ^d	396 (1,7)	OR: 1,49 [1,02; 2,17] ^e
Vergleichsintervention	16 / 29 211 ^d	413 (1,4)	RD: 0,7 [0,0; 1,6] ^f
Chaillet 2015 (Schulung QUARISMA-Programm) ^f			
Prüfintervention	16 / 23 902 ^d	184 (0,8)	OR: 0,98 [0,77; 1,26] ^e
Vergleichsintervention	16 / 29 211 ^d	171 (0,6)	RD: -0,01 [-0,13; 0,15] ^f
Lomas 1991 (Schulung von Meinungsführern) ^g			
Prüfintervention	4 / 345 ^h	7 (0,9)	0,78 [0,32; 1,90] ⁱ
Vergleichsintervention	8 / 410 ^h	15 (1,2)	
Lomas 1991 (Audit und Feedback)			
Prüfintervention	4 / 159 ^h	31 (5,9)	4,86 [2,65; 8,93] ⁱ
Vergleichsintervention	8 / 410 ^h	15 (1,2)	
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Neugeborener c: Apgar-Score nach 5 Minuten ≥ 4 und < 7 d: Anzahl Neugeborener in der Postinterventionsphase e: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) f: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungssatz identisch zu OR g: Apgar-Score nach 5 Minuten < 4 h: Anzahl Neugeborener von Teilnehmerinnen mit vorherigem Kaiserschnitt, bei denen der Versuch einer Spontangeburt laut Leitlinie möglich war i: eigene Berechnung ohne Berücksichtigung der Clusterstruktur KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; RR: relatives Risiko; SSW: Schwangerschaftswoche			

C.1.6 – Schwere metabolische Acidose (pH < 7,0)

Die folgende Tabelle 43 zeigt die Ergebnisse für schwere metabolische Acidose.

Tabelle 43: Ergebnisse für schwere metabolische Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Schwere metabolische Acidose (pH < 7,0)	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 902	87 (0,4)	1,13 [0,79; 1,62] ^c
Vergleichsintervention	16 / 29 211	90 (0,3)	RD: 0,04 [-0,06; 0,19] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Neugeborener in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; OR: adjustiertes Odds Ratio; n: Anzahl; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.7 – Episiotomie

Die folgende Tabelle 44 zeigt die Ergebnisse für Episiotomie.

Tabelle 44: Ergebnisse für Episiotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Episiotomie	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 20 612 ^c	2953 (14,3)	1,01 [0,85; 1,21] ^c
Vergleichsintervention	16 / 24 874 ^c	3871 (15,6)	RD: 0,1 [-2,0; 2,7] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase, bei denen kein geplanter Kaiserschnitt vorgesehen war c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.8 – Wehenmittel

Tabelle 45 und Tabelle 46 zeigen die Ergebnisse für Wehenmittelgabe während der Geburt und zur Einleitung der Geburt.

Tabelle 45: Ergebnisse für Wehenmittel während der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Wehenmittelgabe während der Geburt	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 20 612 ^c	6205 (30,1)	1,16 [1,09; 1,23] ^c
Vergleichsintervention	16 / 24 874 ^c	7572 (30,4)	RD: 3,2 [1,9; 4,6] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase, bei denen kein geplanter Kaiserschnitt vorgesehen war c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungssset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

Tabelle 46: Ergebnisse für Wehenmittel zur Einleitung der Geburt zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Wehenmittelgabe zur Einleitung der Geburt	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 484 ^c	5501 (23,4) ^c	0,82 [0,76; 0,87] ^c
Vergleichsintervention	16 / 28 781 ^c	7872 (27,4) ^c	RD: -3,8 [-5,1; -2,7] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.9 – Amniotomie

Die folgende Tabelle 47 zeigt die Ergebnisse für Amniotomie.

Tabelle 47: Ergebnisse für Amniotomie zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Amniotomie	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 484	11 972 (51,0) ^c	1,02 [0,84; 1,22] ^c
Vergleichsintervention	16 / 28 781	14 534 (50,5) ^c	RD: 0,5 [-4,4; 5,0] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.10 – Infektionen

Die folgende Tabelle 48 zeigt die Ergebnisse für Infektionen. Die folgende Tabelle 49 zeigt die ergänzenden Informationen zum Auftreten von neonatalen Infektionen.

Tabelle 48: Ergebnisse für maternale Infektionen zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Infektionen / Sepsis	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 484	1986 (8,5)	0,95 [0,75; 1,19] ^c
Vergleichsintervention	16 / 28 781	2337 (8,1)	RD: -0,4 [-1,9; 1,4] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Teilnehmerinnen in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und der Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

Tabelle 49: Ergänzende Informationen für neonatale Infektionen zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie (Weiterbildung)	n ^a / n ^b	Infektionen / Sepsis (neonatal)	
		Gebärende mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 902	742 (3,1)	0,84 [0,72; 0,98] ^c
Vergleichsintervention	16 / 29 211	1227 (4,2)	RD: -0,7 [-1,1; -0,1] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Neugeborener in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungsset identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.1.11 – Fortgeschrittene metabolische Acidose

Die folgende Tabelle 50 zeigt die Ergebnisse für fortgeschrittene Acidose.

Tabelle 50: Ergebnisse für die fortgeschrittene Acidose zur Erfahrung / Expertise durch Weiterbildungsmaßnahmen

Studie	n ^a / n ^b	fortgeschrittene metabolische Acidose [pH 7,09–7,0]	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI]
Chaillet 2015 (Schulung QUARISMA-Programm)			
Prüfintervention	16 / 23 902	396 (1,7)	1,49 [1,02; 2,17] ^c
Vergleichsintervention	16 / 29 211	413 (1,4)	RD: 0,7 [0,0; 1,6] ^d
a: Anzahl randomisierter Krankenhäuser b: Anzahl randomisierter Neugeborener in der Postinterventionsphase c: adjustiertes OR, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, adjustiert für prä-spezifizierte potenzielle Risikofaktoren der Geburtsklinik (akademischer Status, Versorgungsstatus) und der Gebärenden (Alter, Parität, vorheriger Kaiserschnitt, Schwangerschaftspathologien, Rauchen während der Schwangerschaft, Reproduktionstechniken) und Feten (SSW, Geburtsgewicht, Lage) d: adjustierte RD, misst den Unterschied der Veränderungen von Prä- zur Postinterventionsperiode zwischen den Interventionsgruppen, Adjustierungssatz identisch zu OR KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; RD: adjustierte Risikodifferenz; SSW: Schwangerschaftswoche			

C.2 – Ergebnisse zu Erfahrung / Expertise durch systemimmanente Faktoren

In diesem Abschnitt folgt eine tabellarische Darstellung der Ergebnisse zu den kritischen Endpunkten für die Suche nach vergleichenden Studien zur Erfahrung / Expertise des geburtshilflichen Fachpersonals, welches über systemimmanente Faktoren beschrieben wird. Aus der Kategorie der kritischen Endpunkte wurden Ergebnisse zum Geburtsmodus (siehe Tabelle 51 und Tabelle 52), zur neonatalen Mortalität (siehe Tabelle 53), zu schwerer metabolischer Acidose (siehe Tabelle 55) berichtet. Zusätzlich werden ergänzende Informationen zum Apgar-Score nach 5 Minuten (Tabelle 54) und zur Verlegung auf die Neugeborenen-Intensivstation (Tabelle 56) berichtet.

Aus der Kategorie der wichtigen Endpunkte wurden Ergebnisse zur Episiotomie (siehe Tabelle 57 und Tabelle 58) und zu Wehenmitteln (siehe Tabelle 59 und Tabelle 60) berichtet.

C.2.1 – Geburtsmodus

Die folgenden Tabellen zeigen die Ergebnisse zum Einfluss der Geburtenzahl pro Jahr (Tabelle 51) und der Versorgungsstufe der Geburtskliniken (Tabelle 52) auf den Geburtsmodus.

Tabelle 51: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr der Geburtsklinik)

Studie (Subgruppe)	n ^a	Spontangeburt		Kaiserschnitt		Vaginal-operative Geburt	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b
Coulm 2012							
Exposition: ≥ 3000	k. A.	1227 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (16,5)	1,47 [1,10; 1,96] ^f / [1,03; 2,09] ^{e, f}
Exposition: 2000–2999	k. A.	2022 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (16,4)	1,42 [1,11; 1,83] ^f / [1,05; 1,93] ^{e, f}
Exposition: 1000–1999	k. A.	2519 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (14,8)	1,27 [1,01; 1,61] ^f / [0,96; 1,69] ^{e, f}
Referenzgruppe: < 1000	k. A.	1354 (k. A.) ^c		k. E. ^d	k. E. ^d	k. A. ^c (12,7)	
Wehberg 2018 (Erstgebärende)							
		Vaginale Geburt		Kaiserschnitt			
Exposition: > 3000	41 689	36 950 (88,6) ^e	1,22 [1,16; 1,30] ^{g, e} / 97,5 %-KI: [1,14; 1,30] ^{g, e}	4739 (11,4)	0,82 [0,77; 0,86] ^g / [0,77; 0,88] ^{g, e}		
Referenzgruppe: 1100–3000	27 970	24 554 (87,8) ^e		3416 (12,2)			
Exposition: 1100–3000	27 970	24 554 (87,8) ^e	1,02 [0,92; 1,13] ^{g, e} / 97,5 %-KI: [0,91; 1,15] ^{g, e}	3416 (12,2)	0,98 [0,88; 1,09] ^{g, e} / [0,86; 1,12] ^{g, e}		
Referenzgruppe: < 1100	4343	3800 (87,5) ^e		543 (12,5)			

(Fortsetzung)

Tabelle 51: Ergebnisse zum Geburtsmodus zur Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr der Geburtsklinik) (Fortsetzung)

a: Anzahl ausgewerteter Teilnehmerinnen
b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise).
c: Bezugsgröße unklar
d: Veröffentlichte Angaben beziehen sich auf alle Geburten mit geringem Risiko für Komplikationen ohne Differenzierung von geplanten Kaiserschnitten und Entscheidung im Geburtsverlauf.
e: eigene Berechnung
f: adjustiert für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad) und neonatale Charakteristika (SSW, Geburtsgewicht)
g: adjustiert für maternale Charakteristika (Alter, BMI, Rauchen, Größe, Diabetes, Bluthochdruck, Eklampsie, Kopf-Becken-Missverhältnis, Plazenta Praevia, vorzeitiger Plazentalösung und vorzeitiger Blasensprung), neonatale Charakteristika (SSW, Geschlecht, Fetal Distress, Makrosomie), geburtsbezogene Charakteristika (Geburtseinleitung) und systembezogene Charakteristika (Geburt am Wochenende oder während der Nacht)
BMI: Body-Mass-Index; k. A.: keine Angaben; k. E.: keine Extraktion; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; SSW: Schwangerschaftswoche

Tabelle 52: Ergebnisse zum Geburtsmodus zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren
(Versorgungsstufe der Geburtskliniken)

Studie	n ^a	Spontangeburt		Kaiserschnitt		Vaginal-operative Geburt	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b	Gebärende mit Ereignis n (%)	OR [95 %-KI] / [98,3 %-KI] ^b
Coulm 2012							
Exposition: Level 3	k. A.	1526 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (15,0)	1,30 [1,01; 1,68] ^e / [0,95; 1,77] ^{e, f}
Exposition: Level 2b	k. A.	1491 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (13,0)	0,98 [0,76; 1,26] ^e / [0,72; 1,33] ^{e, f}
Exposition: Level 2a	k. A.	1868 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (16,1)	1,06 [0,86; 1,32] ^e / [0,82; 1,38] ^{e, f}
Referenzgruppe: Level 1	k. A.	2237 (k. A.) ^c	k. A.	k. E. ^d	k. E. ^d	k. A. ^c (15,9)	
<p>a: Anzahl ausgewerteter Teilnehmerinnen b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise). c: Bezugsgröße unklar d: Veröffentlichte Angaben beziehen sich auf alle Geburten mit geringem Risiko ohne Differenzierung von geplanten Kaiserschnitten und Entscheidung im Geburtsverlauf. e: adjustiert für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad) und neonatale Charakteristika (SSW, Geburtsgewicht) f: eigene Berechnung BMI: Body-Mass-Index; k. A.: keine Angaben; k. E.: keine Extraktion; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; SSW: Schwangerschaftswoche</p>							

C.2.2 – Neonatale Mortalität

Die folgende Tabelle 53 zeigt die Ergebnisse zum Einfluss der Geburtenzahl pro Jahr auf die neonatale Mortalität.

Tabelle 53: Ergebnisse zur neonatalen Mortalität zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Studie (Sensitivitätsanalyse)	n ^a	Neonatale Mortalität	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI] ^b
Finnström 2006^c			
Exposition: ≥ 1000–2499	638 183	k. A.	
Referenzgruppe: 500–999	175 388	k. A.	1,22 [1,09; 1,37] ^{d, e}
Referenzgruppe: < 500	31 745	k. A.	1,19 [0,90; 1,59] ^{d, e}
Exposition: ≥ 2500	239 441	k. A.	1,15 [1,06; 1,25] ^e
Referenzgruppe: 1000–2499	638 183	k. A.	
Finnström 2006 (Auswertung für die Klinik im Einzugsgebiet der Gebärenden)^c			
Exposition: ≥ 1000–2499	641 045	k. A.	
Referenzgruppe: 500–999	162 156	k. A.	0,94 [0,85; 1,04] ^{d, e}
Referenzgruppe: < 500	32 234	k. A.	0,87 [0,71; 1,05] ^{d, e}
Exposition: ≥ 2500	630 559	k. A.	0,94 [0,88; 1,00] ^e
Referenzgruppe: 1000–2499	641 045	k. A.	
Karalis 2016^f			
Universitätskrankenhaus	92 139	219 (0,24 ^d)	
Referenzgruppe: ≥ 2000	84 681	227 (0,27 ^d)	0,87 [0,72; 1,05] ^{d, g}
Referenzgruppe: 1000–1999	51 185	155 (0,30 ^d)	0,76 [0,62; 0,93] ^{d, g}
Referenzgruppe: ≤ 999	39 385	108 (0,27 ^d)	0,84 [0,66; 1,05] ^{d, g}
Karalis 2016^h			
Universitätskrankenhaus	92 139	14 (0,02)	
Referenzgruppe: ≥ 2000	84 681	15 (0,02)	0,85 [0,41; 1,79] ^{d, g}
Referenzgruppe: 1000–1999	51 185	13 (0,03)	0,58 [0,27; 1,23] ^{d, g}
Referenzgruppe: ≤ 999	39 385	12 (0,03)	0,47 [0,22; 1,03] ^{d, g}
a: Anzahl ausgewerteter Neugeborener b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise). c: neonatale Todesfälle 1.–27. Tag d: eigene Berechnung e: adjustiert für Geburtsjahr, maternale Charakteristika (Alter, Parität, Rauchen während der Schwangerschaft, Familienstand, BMI) und fetale Charakteristika (SSW) f: frühe neonatale Todesfälle (Tag 1–7) g: adjustiert für maternale Charakteristika (Alter, Parität) h: intrapartale Todesfälle BMI: Body-Mass-Index; k. A: keine Angaben; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; SSW: Schwangerschaftswoche			

C.2.3 – Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6

Die folgende Tabelle 54 zeigt die ergänzenden Informationen zum Einfluss der Geburtenzahl pro Jahr auf das Auftreten eines geringen Apgar-Score nach 5 Minuten. In 1 Studie lagen ausschließlich Informationen zu Apgar-Score nach 5 Minuten < 4. Diese Studie bezog alle lebend geborenen Einlinge in die Auswertung ein. 1 weitere Studie berichtete das Auftreten eines Apgar-Score nach 5 Minuten < 7. Zusätzlich untersucht eine Sensitivitätsanalyse den Einfluss der Geburtenzahl im Einzugsgebiet der Gebärenden.

Tabelle 54: Ergänzende Informationen zum Apgar-Score nach 5 Minuten < 6 (Apgar-Score nach 5 Minuten < 7 oder < 4 zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr))

Studie (Sensitivitätsanalysen)	n ^a	Apgar-Score nach 5 Minuten	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI] ^b
Finnström 2006^c			
Exposition: ≥ 1000–2499	638 183	k. A.	
Referenzgruppe: 500–999	175 388	k. A.	1,09 [0,98; 1,20] ^{d, e}
Referenzgruppe: < 500	31 745	k. A.	1,33 [1,05; 1,69] ^{d, e}
Exposition: ≥ 2500	239 441	k. A.	1,48 [1,37; 1,60] ^e
Referenzgruppe: 1000–2499	638 183	k. A.	
Finnström 2006 (Auswertung für die Klinik im Einzugsgebiet der Gebärenden)^c			
Exposition: ≥ 1000–2499	641 045	k. A.	
Referenzgruppe: 500–999	162 156	k. A.	1,04 [0,95; 1,16] ^{d, e}
Referenzgruppe: < 500	32 234	k. A.	1,18 [0,96; 1,45] ^{d, e}
Exposition: ≥ 2500	630 559	k. A.	1,16 [1,09; 1,23] ^e
Referenzgruppe: 1000–2499	641 045	k. A.	
Karalis 2016^f			
Universitätskrankenhaus	92 139	1342 (1,5)	
Referenzgruppe: ≥ 2000	84 681	1034 (1,2)	1,19 [1,10; 1,30] ^{d, g}
Referenzgruppe: 1000–1999	51 185	793 (1,5)	0,91 [0,83; 1,00] ^{d, g}
Referenzgruppe: ≤ 999	39 385	600 (1,5)	0,90 [0,82; 1,00] ^{d, g}
a: Anzahl ausgewerteter Neugeborener b: Alle Ergebnisse verglichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise). c: Apgar-Score nach 5 Minuten < 4 d: eigene Berechnung e: adjustiert für Geburtsjahr, maternale Charakteristika (Alter, Parität, Rauchen während der Schwangerschaft, Familienstand, BMI) und neonatale Charakteristika (SSW) f: Apgar-Score nach 5 Minuten < 7 g: adjustiert für maternale Charakteristika (Alter, Parität) BMI: Body-Mass-Index; k. A: keine Angaben; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; SSW: Schwangerschaftswoche			

C.2.4 – Schwere metabolische Acidose (pH < 7,0)

Die folgende Tabelle 55 zeigt die Ergebnisse zum Einfluss der Geburtenzahl pro Jahr auf die schwere metabolische Acidose (pH < 7,0).

Tabelle 55: Ergebnisse zur schweren metabolischen Acidose (pH < 7,0) zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Studie	n ^a	Schwere metabolische Acidose (pH < 7,0)	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI] ^b
Karalis 2016			
Universitätskrankenhaus	92 139	361 (0,39)	
Referenzgruppe: ≥ 2000	84 681	235 (0,28)	1,41 [1,19; 1,67] ^{c, d}
Referenzgruppe: 1000–1999	51 185	179 (0,35)	1,06 [0,88; 1,27] ^{c, d}
Referenzgruppe: ≤ 999	39 385	99 (0,25)	1,43 [1,14; 1,79] ^{c, d}
a: Anzahl ausgewerteter Neugeborener			
b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise).			
c: eigene Berechnung			
d: adjustiert für maternale Charakteristika (Alter, Parität)			
KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio			

C.2.5 – Ergänzende Informationen zur Verlegung auf die Neugeborenen-Intensivstation

Die folgende Tabelle 56 zeigt die ergänzenden Informationen zum Einfluss der Geburtenzahl pro Jahr auf die Verlegung auf die Neugeborenen-Intensivstation. In 1 Studie lagen ausschließlich gemeinsame Informationen zur Verlegung auf die Neugeborenen-Intensivstation, die Beobachtungsstation oder in ein anderes Krankenhaus vor und in 1 weiteren Studie liegen gemeinsame Informationen zur Verlegung auf die Neugeborenen-Intensivstation oder Special-Care Nursery vor.

Tabelle 56: Ergänzende Informationen für die Häufigkeit der Verlegung auf die Neugeborenen-Intensivstation zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Studie	n ^a	Verlegung auf die Neugeborenen-Intensivstation	
		Neugeborene mit Ereignis n (%)	OR [95 %-KI] ^b
Karalis 2016^c			
Universitätskrankenhaus	92 139	8750 (9,5)	
Referenzgruppe: ≥ 2000	84 681	6125 (7,2)	1,35 [1,30; 1,39] ^{d, e}
Referenzgruppe: 1000–1999	51 185	4550 (8,9)	1,05 [1,02; 1,10] ^{d, e}
Referenzgruppe: ≤ 999	39 385	2139 (5,4)	1,75 [1,66; 1,85] ^{d, e}
a: Anzahl ausgewerteter Neugeborener			
b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise).			
c: Aufenthalt auf der Neugeborenen-Intensivstation, Beobachtungsstation oder Transfer in ein anderes Krankenhaus			
d: eigene Berechnung des Kehrwertes aus dem berichteten Effektschätzer			
e: adjustiert für maternale Charakteristika (Alter, Parität)			
KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio			

C.2.6 – Episiotomie

Die folgenden Tabellen zeigen die Ergebnisse zum Einfluss der Geburtenzahl pro Jahr auf die Häufigkeit von Episiotomien in Abhängigkeit von der Geburtenzahl pro Jahr (Tabelle 57) und der Versorgungsstufe (Tabelle 58).

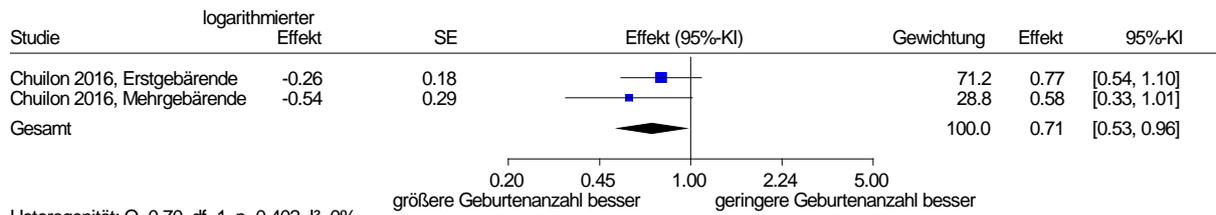
Tabelle 57: Ergebnisse für die Häufigkeit von Episiotomien zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Studie (Subgruppe)	n	Episiotomie	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] ^a
Coulm 2012 (Gebärende mit Spontangeburt)			
Exposition: ≥ 3000	1227 ^b	209 ^c (17,0)	0,80 [0,57; 1,12] ^d
Exposition: 2000–2999	2022 ^b	388 ^c (19,2)	0,95 [0,72; 1,25] ^d
Exposition: 1000–1999	2519 ^b	529 ^c (21,0)	1,13 [0,88; 1,44] ^d
Referenzgruppe: < 1000	1354 ^b	269 ^c (19,9)	
Chuilon 2016 (Erstgebärende mit vaginaler Geburt)			
Exposition: ≥ 4000	219 ^e	91 ^c (41,6)	
Referenzgruppe: 3000–3999	645 ^e	278 ^c (43,1)	RR: 0,93 [0,66; 1,32] ^{f, g}
Referenzgruppe: 2000–2999	1396 ^e	627 ^c (44,9)	RR: 0,96 [0,69; 1,32] ^{f, g}
Referenzgruppe: 1000–1999	1615 ^e	720 ^c (44,6)	RR: 0,88 [0,63; 1,20] ^{f, g}
Referenzgruppe: < 1000	814 ^e	380 ^c (46,7)	RR: 0,77 [0,54; 1,10] ^{f, g}
Chuilon 2016 (Mehrgebärende mit vaginaler Geburt)			
Exposition: ≥ 4000	266 ^e	33 ^c (12,4)	
Referenzgruppe: 3000–3999	865 ^e	106 ^c (12,3)	RR: 0,93 [0,53; 1,61] ^{f, g}
Referenzgruppe: 2000–2999	1852 ^e	272 ^c (14,7)	RR: 0,72 [0,42; 1,25] ^{f, g}
Referenzgruppe: 1000–1999	2270 ^e	343 ^c (15,1)	RR: 0,65 [0,35; 1,18] ^{f, g}
Referenzgruppe: < 1000	1200 ^e	170 ^c (14,2)	RR: 0,58 [0,33; 1,01] ^{f, g}
a: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise).			
b: Anzahl ausgewerteter Teilnehmerinnen mit Spontangeburt			
c: eigene Berechnung			
d: adjustiert für maternale Charakteristika (Alter, Parität, SSW, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad) und neonatale Charakteristika (SSW, Geburtsgewicht)			
e: Anzahl ausgewerteter Teilnehmerinnen mit vaginaler Geburt			
f: Exposition umfasst höhere Erfahrung / Expertise, eigene Berechnung des Kehrwertes aus dem berichteten Effektschätzer			
g: adjustiert für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad, vorhergehender Kaiserschnitt), neonatale Charakteristika (Lage, SSW, Geburtsgewicht) und Charakteristika der Geburt (PDA, Wehenmittelgabe, betreuendes Fachpersonal, Geburtsmodus) bzw. der Geburtsklinik (Größe, Versorgungsstufe, Ausstattung)			
BMI: Body-Mass-Index; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; PDA: Periduralanästhesie; RR: adjustiertes relatives Risiko; SSW: Schwangerschaftswoche			

Tabelle 58: Ergebnisse für die Häufigkeit von Episiotomien zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe)

Studie (Subgruppe)	n	Episiotomie	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] ^a
Coulm 2012 (Gebärende mit Spontangeburt)			
Exposition: Level 3	1526 ^b	253 ^c (16,6)	0,96 [0,70; 1,31] ^d
Exposition: Level 2b	1491 ^b	307 ^c (20,6)	1,29 [0,97; 1,70] ^d
Exposition: Level 2a	1868 ^b	400 ^c (21,4)	1,20 [0,94; 1,53] ^d
Referenzgruppe: Level 1	2237 ^b	436 ^c (19,5)	
Chuilon 2016 (Erstgebärende mit vaginaler Geburt)			
Exposition: Level 3	1048 ^e	425 ^c (40,6)	
Referenzgruppe: Level 2b	954 ^e	443 ^c (46,4)	RR: 0,93 [0,74; 1,18] ^{f, g}
Referenzgruppe: Level 2a	1270 ^e	578 ^c (45,5)	RR: 1,01 [0,87; 1,18] ^{f, g}
Referenzgruppe: Level 1	1414 ^e	650 ^c (46,0)	RR: 1,10 [0,85; 1,41] ^{f, g}
Chuilon 2016 (Mehrgebärende mit vaginaler Geburt)			
Exposition: Level 3	1415 ^e	175 ^c (12,4)	
Referenzgruppe: Level 2b	1375 ^e	208 ^c (15,1)	RR: 0,91 [0,68; 1,22] ^{f, g}
Referenzgruppe: Level 2a	1713 ^e	264 ^c (15,4)	RR: 0,98 [0,71; 1,37] ^{f, g}
Referenzgruppe: Level 1	1944 ^e	278 ^c (14,3)	RR: 1,14 [0,81; 1,59] ^{f, g}
a: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise).			
b: Anzahl ausgewerteter Teilnehmerinnen mit Spontangeburt			
c: eigene Berechnung			
d: adjustiert für maternale Charakteristika (Alter, Parität, SSW, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad) und neonatale Charakteristika (SSW, Geburtsgewicht)			
e: Anzahl ausgewerteter Teilnehmerinnen mit vaginaler Geburt			
f: Exposition umfasst höhere Erfahrung / Expertise, eigene Berechnung des Kehrwertes aus dem berichteten Effektschätzer			
g: adjustiert für maternale (Alter, Parität, BMI vor der Schwangerschaft, Geburtsland, Bildungsgrad, vorhergehender Kaiserschnitt), neonatale Charakteristika (Lage, SSW, Geburtsgewicht) und Charakteristika der Geburt (PDA, Wehenmittelgabe, betreuendes Fachpersonal, Geburtsmodus) bzw. der Geburtsklinik (Größe, Versorgungsstufe, Ausstattung)			
BMI: Body-Mass-Index; KI: Konfidenzintervall; n: Anzahl; OR: adjustiertes Odds Ratio; PDA: Periduralanästhesie; RR: adjustiertes relatives Risiko; SSW: Schwangerschaftswoche			

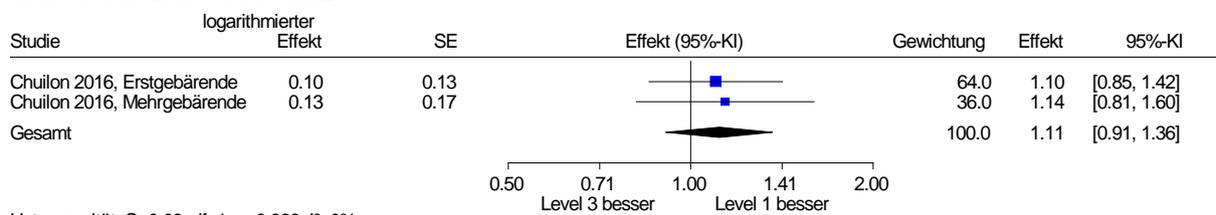
größere Geburtenanzahl vs. geringere Geburtenanzahl
Episiotomie
Modell mit festem Effekt - Inverse Varianz



Effekt entspricht dem OR

Abbildung 4: Forest Plot für Episiotomie zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Level 3 vs. Level 1
Episiotomie
Modell mit festem Effekt - Inverse Varianz



Effekt entspricht dem OR

Abbildung 5: Forest Plot für Episiotomie zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Versorgungsstufe)

C.2.7 – Wehenmittel

Die folgenden Tabellen zeigen die Ergebnisse zum Einfluss der Geburtenzahl pro Jahr auf die Häufigkeit von Wehenmittelgaben während der Geburt und Wehenmittelgaben zur Einleitung der Geburt. 1 Studie beschreibt die Häufigkeit von Wehenmittelgaben während der Geburt oder zur Geburtseinleitung in Abhängigkeit von der Geburtenzahl pro Jahr (Tabelle 59) und 1 Studie beschreibt die Häufigkeit der Wehenmittelgaben zur Einleitung der Geburt in Abhängigkeit von der Versorgungsstufe (Tabelle 60).

Tabelle 59: Ergebnisse für die Häufigkeit von Wehenmittelgaben während der Geburt zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren (Geburtenzahl pro Jahr)

Studie (Subgruppe)	n ^a	Wehenmittelgabe	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] ^b
Belghiti 2013			
Exposition: ≥ 4000	256	159 ^c (62,0)	
Referenzgruppe: 3000–3999	797	542 ^c (68,0)	0,71 [0,53; 1,0] ^{c, d}
Referenzgruppe: 2000–2999	1745	1239 ^c (71,0)	0,83 [0,71; 1,0] ^{c, d}
Referenzgruppe: 1000–1999	1956	1428 ^c (73,0)	0,77 [0,67; 0,91] ^{c, d}
Referenzgruppe: < 1000	914	695 ^c (76,0)	0,67 [0,56; 0,83] ^{c, d}
a: Anzahl ausgewerteter Teilnehmerinnen mit geringem Risiko für Komplikationen und französischer Standardbetreuung im Jahr 2010 (spontan eingesetzte Wehentätigkeit und PDA) b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise). c: eigene Berechnung d: adjustiert für maternale Charakteristika (Alter, Parität, BMI vor der Schwangerschaft) BMI: Body-Mass-Index; KI: Konfidenzintervall; n: Anzahl der Gebärenden mit Ereignis; OR: adjustiertes Odds Ratio; PDA: Periduralanästhesie			

Tabelle 60: Ergebnisse für die Häufigkeit der Wehenmittelgaben zur Einleitung der Geburt zur Beschreibung von Erfahrung / Expertise durch systemimmanente Faktoren

Studie (Subgruppe)	n ^a	Wehenmittelgabe	
		Gebärende mit Ereignis n (%)	OR [95 %-KI] ^b
Belghiti 2013			
Exposition: Level 3	1261	832 ^c (66,0)	
Referenzgruppe: Level 2b	1175	811 ^c (69,0)	0,83 [0,71; 1,0] ^{c, d}
Referenzgruppe: Level 2a	1565	1142 ^c (73,0)	0,71 [0,59; 0,83] ^{c, d}
Referenzgruppe: Level 1	1659	1244 ^c (75,0)	0,63 [0,51; 0,77] ^{c, d, e}
a: Anzahl ausgewerteter Teilnehmerinnen mit geringem Risiko für Komplikationen und französischer Standardbetreuung im Jahr 2010 (spontan eingesetzte Wehentätigkeit und PDA) b: Alle Ergebnisse vergleichen die Expositionsgruppen (größere Erfahrung / Expertise beschrieben durch eine höhere Geburtenzahl pro Jahr) mit den Referenzgruppen (geringere Erfahrung / Expertise). c: eigene Berechnung d: adjustiert für maternale Charakteristika (Alter, Parität, BMI vor der Schwangerschaft) e: Berichtet wurde OR 1,6 (95 %-KI: 1,3–1,4), Korrektur der oberen Grenze des 95 %-KI, da der Effektschätzer nicht im KI lag über die Berechnung des Standardfehlers von lnOR. BMI: Body-Mass-Index; KI: Konfidenzintervall; n: Anzahl der Gebärenden mit Ereignis; OR: adjustiertes Odds Ratio; PDA: Periduralanästhesie			