

Aktualisierungsrecherche zum Bericht S07-01 – Screening auf Gestationsdiabetes

Arbeitspapier

Version 1.0
Stand: 25.03.2010

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Aktualisierungsrecherche zum Bericht S07-01 – Screening auf Gestationsdiabetes

Auftraggeber des Auftrags S07-01:

Gemeinsamer Bundesausschuss

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Dillenburger Str. 27
51105 Köln

Tel.: 0221/35685-0

Fax: 0221/35685-1

Berichte@iqwig.de

www.iqwig.de

Dieses Arbeitspapier wurde einem externen Review unterzogen. Das Review wurde durchgeführt von:

- Dr. Karl Horvath
EbM Review Center, Medizinische Universität Graz

Die Offenlegung der potenziellen Interessenkonflikte des externen Reviewers findet sich in Anhang C.

Kurzfassung

Fragestellung

Unter Gestationsdiabetes (GDM, für: gestational diabetes mellitus) wird im Allgemeinen jede erstmalig im Rahmen einer Schwangerschaft auftretende oder diagnostizierte Einschränkung der Glukosetoleranz (IGT, für: impaired glucose tolerance) verstanden [1,2]. Diese Glukosestoffwechselstörung kann in unterschiedlichen Schweregraden auftreten, die von einer gering eingeschränkten Glukosetoleranz bis zu einem manifesten Diabetes mellitus reichen.

Die Bedeutung der Diagnose Gestationsdiabetes, der Zusammenhang mit Schwangerschafts- und Geburtsrisiken und die bestehenden Unsicherheiten sind im Abschlussbericht „Screening auf Gestationsdiabetes“ (Auftrag S07-01) ausführlich dargestellt [3]. Im Laufe der Bearbeitung des Auftrags wurde zu Teilziel 2 „Therapien“ eine laufende Studie (Landon et al. [4,5]) identifiziert, die die realistische Möglichkeit eröffnete, bestehende Unsicherheiten im Abschlussbericht zu verringern. Die Ergebnisse dieser Studie wurden bis zur Publikation des Abschlussberichts nur als Abstract und auf Kongressen präsentiert, waren aber nicht abschließend als Vollpublikation verfügbar. Die Studie konnte deshalb bis zum Juli 2009 nicht eingeschlossen werden.

Mittlerweile wurde Landon et al. publiziert [6]. Dieses Arbeitspapier beschreibt die Auswirkungen der Studie auf die Ergebnisse und Schlussfolgerungen des Abschlussberichts S07-01. Ziele sind:

- die Aktualisierung der im Abschlussbericht S07-01 beschriebenen Literaturrecherchen zu den Teilzielen 1 und 2. Im Vergleich zum Abschlussbericht wurde die Suche auf weniger Datenbanken (MEDLINE, EMBASE, Cochrane) begrenzt und für Teilziel 2 auf randomisierte Studien eingeschränkt, die sich Studienpool A zuordnen ließen.
- die Extraktion der Ergebnisse der identifizierten Studien
- eine Abschätzung der Auswirkungen auf die Schlussfolgerungen des Abschlussberichts S07-01

Methoden

Für die vorliegende Untersuchung wurde systematisch nach Studien für die Teilfragestellung gesucht. Die Ergebnisse zu patientenrelevanten Endpunkten und definierten Surrogatparametern wurden extrahiert und zusammenfassend bewertet.

Die ausführlichen Ziele des Projekts S07-01, Methoden, der Projektablauf und die Ergebnisse sind im Abschlussbericht beschrieben, der im August 2009 publiziert wurde [3]. Für Teilziel 1 „Screening“ kamen wie im Abschlussbericht in erster Linie randomisierte kontrollierte Studien (RCTs) an unselektierten Schwangeren infrage. Nicht-randomisierte kontrollierte

Studien (nRCTs) konnten herangezogen werden, sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) adäquat adressiert wurde und zwischen den Kollektiven vergleichbare Nebenbedingungen vorlagen. Für Teilziel 2 „Therapie“ kamen im Unterschied zum Abschlussbericht nur RCTs infrage, in denen eine Intervention mit blutzuckersenkender Intention oder eine anderweitige Intervention, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet ist, verglichen wurde mit einem Vorgehen ohne eine solche Intervention („Routinebetreuung“). Der Studienpool zum Vergleich von Therapien unterschiedlicher Intensität wurde nicht aktualisiert. Diese Einschränkungen basieren darauf, dass die Ergebnisse des Studienpools A bereits auf RCTs beruhten und nur weitere RCTs das Potenzial haben, diese Ergebnisse zu modifizieren.

Für das vorliegende Arbeitspapier wurden in Übereinstimmung mit den Methoden des Projekts S07-01 insbesondere nachfolgende Zielgrößen verwendet, die eine Bewertung patientenrelevanter Therapieziele ermöglichten. Mütterliche Ergebnisse: Mortalität, Art der Geburt, Geburtskomplikationen (zum Beispiel Schulterdystokie), Präeklampsie / Eklampsie. Kindliche Ergebnisse: perinatale und neonatale Mortalität, Geburtstraumata, diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen, Aufnahme auf eine Intensivstation, unerwünschte Ereignisse. Makrosomie / Geburtsgewicht war kein patientenrelevanter Endpunkt, sondern ein Surrogatparameter unklarer Validität.

Ergebnisse

Die Kernfrage des Auftrags S07-01 konnte auch durch dieses Arbeitspapier nicht abschließend beantwortet werden. In Teilziel 1 „Screening“ konnten weiterhin keine Studien gefunden werden, in denen ein Screening auf Gestationsdiabetes mit keinem Screening verglichen wurde. Die Gesamtbewertung für Teilziel 1 lautete deshalb weiterhin, dass für keinen patientenrelevanten Endpunkt ein Beleg für einen Nutzen oder Schaden durch ein Screening auf Gestationsdiabetes vorlag.

Zur Bewertung des Nutzens von Therapien (Teilziel 2) konnte eine neue Studie in Pool A aufgenommen werden (Landon 2009). Aus den 5 RCTs des aktualisierten Studienpools A ergibt sich folgendes Bild zum Vergleich zwischen GDM-spezifischen Therapien und einer Routinebetreuung: In der Gesamtabwägung wird Crowther 2005 weiterhin als Hinweis auf einen Nutzen hinsichtlich „ernster perinataler Komplikationen“ bewertet. Die Größe des Nutzens bleibt jedoch unklar. Für den Endpunkt Schulterdystokie liegt jetzt ein Beleg für einen Nutzen durch eine GDM-spezifische Therapie vor. Für den Endpunkt Präeklampsie liegt jetzt ein Hinweis auf einen Nutzen vor. Dieser Hinweis basiert auf Landon 2009.

Meta-Analysen ergaben einen statistisch signifikant geringeren Anteil makrosomer / Large-for-Gestational-Age-Kinder durch eine GDM-spezifische Therapie.

Schäden durch eine GDM-spezifische Therapie wurden in den Studien nicht explizit untersucht und infolgedessen auch nicht berichtet. Eine Meta-Analyse für Kinder „Small-for-Gestational-Age“ ergab keinen statistisch signifikanten Unterschied.

Auch wenn ein Hinweis auf einen Nutzen einer GDM-spezifischen Therapie vorliegt, ergibt sich daraus nicht zwangsläufig auch ein Hinweis auf einen Nutzen durch ein Screening. Als Aspekte eines potenziellen Schadens wurden betrachtet: Aufwand für den Test, unerwünschte Ereignisse des oGTT, nachteilige psychologische Effekte, falsch-negative Testergebnisse und Effekte durch Risikokompensation. Zusammenfassend ergab sich aus diesen Überlegungen zwar das Potenzial für Schäden durch ein Screening auf Gestationsdiabetes, die potenziellen Risiken durch ein Screening für Frauen mit negativem oGTT wurden aber nicht als schwerwiegend bewertet.

Eine direkte Aussage über Nutzen und Schäden des Screenings war weiterhin nicht möglich. Es wurde aber ein Beleg für einen Nutzen einer GDM-spezifischen Therapie hinsichtlich der Reduktion von Schulterdystokien gefunden, für den breiter gefassten kombinierten Endpunkt perinatale Komplikationen blieb es bei einem Hinweis. Indirekt wurde daraus ein Hinweis abgeleitet, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt. Diese Ableitung basiert auf der Annahme, dass ein Screening eine Population identifiziert, wie sie in den maßgeblichen Therapiestudien eingeschlossen wurde.

Abschätzung der Auswirkungen auf das Fazit des Abschlussberichtes S07-01

Das Fazit des Abschlussberichtes S07-01 lautete:

„Es liegt ein Hinweis auf einen Nutzen einer Gestationsdiabetes-spezifischen Therapie vor.

Es existieren keine direkten Belege für oder Hinweise auf einen Nutzen oder einen Schaden durch ein Screening auf Gestationsdiabetes, da keine geeigneten Screeningstudien identifiziert wurden.

Trotzdem kann indirekt ein Hinweis darauf abgeleitet werden, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt.“

Die Ergebnisse dieses Arbeitspapiers würden zu folgender Änderung des Fazits des Abschlussberichtes S07-01 führen (Änderungen kursiv):

Es liegt ein *Beleg* für den Nutzen einer Gestationsdiabetes-spezifischen Therapie vor.

Es existieren keine direkten Belege für oder Hinweise auf einen Nutzen oder einen Schaden durch ein Screening auf Gestationsdiabetes, da keine geeigneten Screeningstudien identifiziert wurden.

Trotzdem kann indirekt ein Hinweis darauf abgeleitet werden, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt.

Schlagwörter: eingeschränkte Glukosetoleranz, Schwangerschaftsdiabetes, Gestationsdiabetes, Screening, gestational diabetes, IGT, GIGT, oGTT, Glukosebelastungstest, Glukosetoleranztest, systematische Übersicht

Inhaltsverzeichnis

	Seite
Kurzfassung	iii
Tabellenverzeichnis	x
Abbildungsverzeichnis	xi
Abkürzungsverzeichnis	xii
1 Hintergrund	1
2 Ziele der Untersuchung	2
2.1 Teilziel 1 „Screening“	2
2.2 Teilziel 2 „Therapie“	2
3 Projektablauf	3
4 Methoden	4
4.1 Kriterien für den Einschluss von Studien in die Untersuchung	4
4.1.1 Population.....	4
4.1.2 Intervention und Vergleichsbehandlung	5
4.1.3 Zielgrößen	5
4.1.4 Surrogatparameter	6
4.1.5 Studientypen.....	7
4.1.6 Ein- / Ausschlusskriterien	7
4.2 Informationsbeschaffung	9
4.2.1 Literaturrecherche	9
4.2.2 Identifizierung relevanter Studien.....	9
4.2.3 Suche nach zusätzlichen Informationen zu relevanten Studien	10
4.3 Informationsbewertung	10
4.4 Informationssynthese und -analyse	12
4.4.1 Charakterisierung der Studien.....	12
4.4.2 Meta-Analyse	12
4.4.3 Sensitivitätsanalyse	13
4.4.4 Subgruppenanalyse.....	13
5 Ergebnisse	14
5.1 Teilziel 1 „Screening“	14

5.1.1	Ergebnis der Literaturrecherche	14
5.2	Teilziel 2 „Therapie“	15
5.2.1	Ergebnis der Literaturrecherche	15
5.2.1.1	Resultierender Studienpool	16
5.2.1.2	Studienpool A: Charakteristika der in die Bewertung eingeflossenen randomisierten Studien.....	17
5.2.1.2.1	Studiendesign und Studienpopulationen	17
5.2.1.3	Studienpool A: Verzerrungspotenzial der Studien.....	25
5.2.1.4	Studienpool A: Ergebnisse der randomisierten Studien zu patientenrelevanten Therapiezielen.....	29
5.2.1.4.1	Mütterliche Mortalität	29
5.2.1.4.2	Art der Geburt	29
5.2.1.4.3	Andere Geburtsinterventionen	30
5.2.1.4.4	Geburtsverletzungen.....	30
5.2.1.4.5	Geburtskomplikationen (zum Beispiel Schulterdystokie, postpartale Blutungen)	31
5.2.1.4.6	Präeklampsie / Eklampsie	32
5.2.1.4.7	Dauer des stationären Aufenthalts.....	32
5.2.1.4.8	Inanspruchnahme medizinischer Leistungen (Untersuchungen, Tests, Behandlungen)	33
5.2.1.4.9	Psychologische Ergebnisse (zum Beispiel Angst, postnatale Depression).....	33
5.2.1.4.10	Gesundheitsbezogene Lebensqualität.....	33
5.2.1.4.11	Perinatale und neonatale Mortalität.....	34
5.2.1.4.12	Anteil der Fehlgeburten.....	35
5.2.1.4.13	Geburtstraumata (zum Beispiel Verletzungen des Brachialplexus, Knochenbrüche)	35
5.2.1.4.14	Diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen (zum Beispiel Glukoseinfusionen aufgrund von Hypoglykämie, Fototherapie aufgrund von Hyperbilirubinämie, Beatmung aufgrund eines akuten Atemnotsyndroms) inkl. Aufnahme auf eine Intensivstation und Dauer des stationären Aufenthalts	36
5.2.1.4.15	Unerwünschte Folgen einer Intervention mit blutzuckersenkender Intention oder anderer Interventionen, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet sind	39
5.2.1.4.16	Geburtsgewicht / Makrosomie / LGA / SGA	40
5.2.1.4.17	Apgar-Wert.....	42

5.2.1.4.18	Schwangerschaftsbedingte Hypertonie	43
5.2.1.4.19	Spätere Diagnose eines Diabetes mellitus der Mutter.....	45
5.2.1.4.20	Inanspruchnahme medizinischer Leistungen (Untersuchungen, Tests, Behandlungen)	45
5.2.1.4.21	Einleitung der Geburt	46
5.2.1.4.22	Ergebnisse zu Laborparametern	47
5.2.1.4.23	Angaben zum Gestationsalter bei Geburt und Frühgeburtlichkeit.....	48
5.2.1.4.24	Spätere Entwicklung von Übergewicht des Kindes	49
5.2.1.5	Landon 2009: Bewertung des kombinierten Endpunkts	50
5.2.1.6	Studienpool A: Zusammenfassung der Ergebnisse.....	50
5.3	Zusammenfassung der Änderungen aufgrund der Aktualisierung	52
6	Diskussion	53
7	Abschätzung der Auswirkungen auf das Fazit des Abschlussberichtes S07-01	64
8	Liste der eingeschlossenen Studien.....	65
8.1	Liste der durch die Nachrecherche eingeschlossenen Studien.....	65
8.2	Liste der bereits im Abschlussbericht S07-01 eingeschlossenen Studien zum Teilziel 2 „Therapie“, Studienpool A	65
9	Literatur.....	67
Anhang A: Suchstrategien.....		71
Anhang B: Liste der im Volltext überprüften, aber ausgeschlossenen Studien mit Ausschlussgründen.....		74
Anhang C: Offenlegung potenzieller Interessenkonflikte des externen Reviewers		75

Tabellenverzeichnis

Tabelle 1: Einschlusskriterien Teilziel 1 (S = Screening).....	8
Tabelle 2: Einschlusskriterien Teilziel 2 (T = Therapie)	8
Tabelle 3: Ausschlusskriterien (alle Teilziele betreffend)	9
Tabelle 4: Quellen der Information.....	9
Tabelle 5: Studienpool A – Ergebnisse der Literaturbeschaffung	17
Tabelle 6: Studienpool A – Übersicht über die eingeschlossenen randomisierten Studien	20
Tabelle 7: Studienpool A – Basischarakteristika der Studienpopulationen	22
Tabelle 8: Studienpool A – Beschreibung der Interventionen und der Ein- und Ausschlusskriterien.....	23
Tabelle 9: Studienpool A – Einschätzung des Verzerrungspotenzials der eingeschlossenen Studien	27
Tabelle 10: Studienpool A – Kaiserschnittraten	30
Tabelle 11: Studienpool A – Schulterdystokie.....	31
Tabelle 12: Studienpool A – perinatale und neonatale Mortalität	35
Tabelle 13: Studienpool A – Geburtstraumata	36
Tabelle 14: Studienpool A – diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen	39
Tabelle 15: Studienpool A – Geburtsgewicht und verwandte Endpunkte	42
Tabelle 16: Studienpool A – Apgar-Werte	43
Tabelle 17: Studienpool A – Schwangerschaftsbedingte Hypertonie	44
Tabelle 18: Crowther 2005 – Inanspruchnahme medizinischer Leistungen nach Aufnahme in die Studie (nur Einlingsschwangerschaften)	46
Tabelle 19: Studienpool A – Einleitung der Geburt.....	47
Tabelle 20: Studienpool A – Ergebnisse zu Laborparametern.....	48
Tabelle 21: Studienpool A – Gestationsalter bei Geburt und Frühgeburtenrate	49
Tabelle 22: Landon 2009 – Ergebnisse des primären kombinierten Endpunkts.....	50
Tabelle 23: HAPO 2008 – Konsensusvorschlag für Grenzwerte eines Screenings auf Gestationsdiabetes	59
Tabelle 24: HAPO 2008 – Assoziationen zwischen BG und Risiken für ausgewählte Endpunkte.....	61

Abbildungsverzeichnis

Abbildung 1: Teilziele 1 und 2 – Ergebnisse der Aktualisierungsrecherche	16
Abbildung 2: Studienpool A – Meta-Analyse Anteil der Geburten mit Kaiserschnitt	29
Abbildung 3: Studienpool A – Meta-Analyse der Schulterdystokien.....	31
Abbildung 4: Studienpool A – Forest-Plot perinatale und neonatale Mortalität.....	34
Abbildung 5: Studienpool A – Meta-Analyse der Geburtstraumata	36
Abbildung 6: Studienpool A – Meta-Analyse der Glukoseinfusionen	37
Abbildung 7: Studienpool A – Meta-Analyse Aufnahme auf eine Intensivstation.....	38
Abbildung 8: Studienpool A – Meta-Analyse Makrosomie.....	40
Abbildung 9: Studienpool A – Meta-Analyse Large-for-Gestational-Age.....	41
Abbildung 10: Studienpool A – Meta-Analyse Small-for-Gestational-Age.....	41
Abbildung 11: Studienpool A – Meta-Analyse schwangerschaftsbedingte Hypertonie.....	44
Abbildung 12: Studienpool A – Meta-Analyse Einleitung der Geburt.....	47

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ADA	American Diabetes Association
AGMFM	Arbeitsgemeinschaft für materno-fetale Medizin
BG	Blutglukosekonzentration
BMI	Body Mass Index
BQS	BQS Bundesgeschäftsstelle Qualitätssicherung GgmbH
CONSORT	Consolidated Standards of Reporting Trials
DGGG	Deutsche Gesellschaft für Gynäkologie und Geburtshilfe
G-BA	Gemeinsamer Bundesausschuss
GDM	Gestationsdiabetes (gestational diabetes mellitus)
GCT	Glukose-Challenge-Test
GIGT	Gestational Impaired Glucose Tolerance
HTA	Health Technology Assessment
HAPO	Hyperglycemia and Adverse Pregnancy Outcome
IG	Interventionsgruppe
IGT	eingeschränkte Glukosetoleranz (impaired glucose tolerance)
IQR	Interquartilsabstand (interquartile range)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to Treat
KG	Kontrollgruppe
KI	Konfidenzintervall
LGA	Large-for-Gestational-Age
MDI	Multiple Daily Injections
NDDG	National Diabetes Data Group
NPH	Neutrales Protamin Hagedorn
nRCT	Non-randomised Controlled Trial
NüBG	Nüchtern-Blutglukose-Konzentration
oGTT	oraler Glukosetoleranztest
OR	Odds Ratio
RCT	Randomised Controlled Trial
RF	Risikofaktor

Abkürzung	Bedeutung
SD	Standardabweichung (standard deviation)
SGA	Small-for-Gestational-Age
SOGC	Society of Obstetricians and Gynaecologists of Canada
SSW	Schwangerschaftswoche
STARD	Standards for Reporting Studies of Diagnostic Accuracy
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Nonrandomised Designs
WHO	Weltgesundheitsorganisation (World Health Organization)

1 Hintergrund

Unter Gestationsdiabetes (GDM, für: gestational diabetes mellitus) wird im Allgemeinen jede erstmalig im Rahmen einer Schwangerschaft auftretende oder diagnostizierte Einschränkung der Glukosetoleranz (IGT, für: impaired glucose tolerance) verstanden [1,2]. Diese Glukosestoffwechselstörung kann in unterschiedlichen Schweregraden auftreten, die von einer gering eingeschränkten Glukosetoleranz bis zu einem manifesten Diabetes mellitus reichen.

Die Bedeutung der Diagnose Gestationsdiabetes, der Zusammenhang mit Schwangerschafts- und Geburtsrisiken und die bestehenden Unsicherheiten sind im IQWiG-Abschlussbericht „Screening auf Gestationsdiabetes“ (Auftrag S07-01, veröffentlicht im August 2009) ausführlich dargestellt [3].

Hauptziel (1) des Auftrags war die Nutzenbewertung eines Screenings auf eingeschränkte Glukosetoleranz in der Schwangerschaft hinsichtlich patientenrelevanter Therapieziele. Weitere Teilziele waren die Bewertung (2) der Effektivität von Therapien, (3) des Zusammenhangs zwischen den Ergebnissen eines oralen Glukosetoleranztests (oGTT) und mütterlichen wie kindlichen Outcomes und (4) der Vergleichbarkeit der in den Studien zu diesen letzten beiden Teilzielen eingeschlossenen Frauen.

Im Laufe der Bearbeitung des Auftrags wurde zu Teilziel 2 „Therapien“ eine laufende Studie (Landon et al. [4,5]) identifiziert, die die realistische Möglichkeit eröffnete, bestehende Unsicherheiten im Abschlussbericht zu verringern. Die Ergebnisse dieser Studie wurden bis zur Publikation des Abschlussberichts nur als Abstract und auf 2 Kongressen präsentiert, waren aber nicht abschließend als Vollpublikation verfügbar. Die Studie konnte deshalb bis zum Juli 2009 nicht eingeschlossen werden.

Mittlerweile wurden die Studienergebnisse als Vollpublikation Landon et al. [6] veröffentlicht. Dieses Arbeitspapier aktualisiert die Recherche für die Teilziele 1 und 2 und beschreibt die Auswirkungen von Landon et al. [6] auf die Ergebnisse und Schlussfolgerungen des Abschlussberichts S07-01.

2 Ziele der Untersuchung

Ziele dieses Arbeitspapiers sind:

- die Aktualisierung der im Abschlussbericht S07-01 beschriebenen Ergebnisse zu den Teilzielen 1 und 2 auf Basis einer angepassten, enger gefassten Recherche
- die Extraktion der Ergebnisse der identifizierten Studien
- eine Abschätzung der Auswirkungen auf die Schlussfolgerungen des Abschlussberichts S07-01.

Die ausführlichen Ziele des Projekts S07-01, Methoden, der Projektablauf und die Ergebnisse sind im Abschlussbericht [3] beschrieben. Für das vorliegende Arbeitspapier wurde Teilziel 1 unverändert beibehalten. Teilziel 2 wurde eingeschränkt auf Vergleiche einer GDM-spezifischen Therapie mit einer Routinebehandlung.

2.1 Teilziel 1 „Screening“

Bewertung der Effektivität von Screeningprogrammen:

- vergleichende Nutzenbewertung eines Screenings auf Gestationsdiabetes unter Einsatz eines oGTT mit einem Vorgehen ohne Screening hinsichtlich patientenrelevanter Endpunkte
- vergleichende Nutzenbewertung eines Screenings auf Gestationsdiabetes unter Einsatz eines oGTT mit anderen auf blutzuckerassoziierte Risiken zielenden Screeningstrategien (zum Beispiel Vergleich verschiedener oraler Glukosetoleranztests, universelles vs. strategisches Screening) hinsichtlich patientenrelevanter Endpunkte

2.2 Teilziel 2 „Therapie“

Bewertung der Effektivität einer Therapie:

- Bewertung einer Intervention mit blutzuckersenkender Intention oder einer anderweitigen Intervention, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet ist, bei Frauen, bei denen in einem Screeningkontext (unselektierte Frauen) ein Gestationsdiabetes diagnostiziert wurde, jeweils im Vergleich zu einem Vorgehen ohne eine solche Intervention

3 Projektablauf

Der Verlauf des Projekts S07-01 ist im Abschlussbericht [3] beschrieben. Für dieses Arbeitspapier wurde im Oktober 2009 eine interne Projektskizze erstellt, in der das methodische Vorgehen vorab beschrieben wurde.

4 Methoden

Die Methoden des vorliegenden Arbeitspapiers entsprechen denen des Abschlussberichts [3] mit den folgenden Anpassungen:

- Die Nachrecherche zielte nur auf die Teilziele 1 und 2 ab.
- Die Nachrecherche wurde auf die Datenbanken CENTRAL, EMBASE und MEDLINE eingeschränkt, da alle für den Abschlussbericht S07-01 relevanten Studien in diesen Datenbanken identifiziert worden waren.
- Der Einschluss von Studien zu Teilziel 2 wurde eingeschränkt auf randomisierte kontrollierte Studien (RCTs), die dem Studienpool A des Abschlussberichts zugeordnet werden können (siehe 4.1.2; 4.1.5). Diese Einschränkungen basieren darauf, dass die Ergebnisse des Studienpools A bereits auf RCTs beruhten und nur weitere RCTs das Potenzial haben, diese Ergebnisse zu modifizieren.

Studienpool A enthält Studien, in denen eine GDM-spezifische Therapie mit keiner Therapie oder keiner GDM-spezifischen Therapie verglichen wurde. Diese Festlegung auf Studien zu Pool A hat zur Folge, dass in diesem Arbeitspapier Studien, in denen GDM-spezifische Therapien unterschiedlicher Intensität verglichen wurden, ausgeschlossen wurden. Diese Studien waren im Abschlussbericht Studienpool B zugeordnet worden. Da die bereits im Abschlussbericht S07-01 beschriebene Heterogenität der Interventionen in diesem Studienpool die Ergebnisse schwer interpretierbar macht, wurde auf eine Aktualisierung dieses Studienpools verzichtet.

Im Folgenden sind die Methoden mit diesen Anpassungen beschrieben.

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

Nachfolgend sind die Kriterien genannt, die Voraussetzung für den Einschluss einer Studie in die Nutzenbewertung waren (Einschlusskriterien) oder zu einem Ausschluss aus der weiteren Bewertung führten (Ausschlusskriterien).

4.1.1 Population

Teilziel 1 „Screening“:

unselektierte Schwangere

Teilziel 2 „Therapie“:

Schwangere, bei denen in einem Screeningsetting eine Einschränkung der Glukosetoleranz diagnostiziert wurde

4.1.2 Intervention und Vergleichsbehandlung

Aus den in Kapitel 2 definierten Teilzielen ergaben sich die folgenden Vergleiche:

Teilziel 1 „Screening“:

Verglichen wurde die Durchführung eines Screenings auf Gestationsdiabetes entweder mit der Nichtdurchführung eines Screenings oder mit der Durchführung einer anderen Screeningstrategie wie unter 2.1 definiert.

Teilziel 2 „Therapie“:

Verglichen wurden Interventionen mit blutzuckersenkender Intention oder Interventionen, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet sind, mit einem Vorgehen ohne eine solche Intervention. In Interventions- und Kontrollgruppen mussten ansonsten vergleichbare Nebenbedingungen vorliegen.

4.1.3 Zielgrößen

Für dieses Arbeitspapier wurden dieselben patientenrelevanten Zielgrößen wie im Abschlussbericht S07-01 verwendet. Das sind:

Mütterliche Ergebnisse:

- Mortalität
- Art der Geburt (zum Beispiel operative Entbindung)
- andere Geburtsinterventionen (zum Beispiel Dammschnitt, Analgesie)
- Geburtsverletzungen (zum Beispiel Dammriss Grad 3 / 4, Verletzungen des Urogenitaltraktes, Uterusruptur)
- Geburtskomplikationen (zum Beispiel Schulterdystokie, postpartale Blutungen)
- Präeklampsie / Eklampsie
- Dauer des stationären Aufenthalts
- Inanspruchnahme medizinischer Leistungen (Untersuchungen, Tests, Behandlungen)
- psychologische Ergebnisse (zum Beispiel postnatale Depression, Angst)

Kindliche Ergebnisse:

- perinatale und neonatale Mortalität

- Anteil der Fehlgeburten
- Geburtstraumata (zum Beispiel Verletzungen des Brachialplexus, Knochenbrüche)
- diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen (zum Beispiel Glukoseinfusionen aufgrund von Hypoglykämie, Fototherapie aufgrund von Hyperbilirubinämie, Beatmung aufgrund eines akuten Atemnotsyndroms)
- Aufnahme auf eine Intensivstation
- Dauer des stationären Aufenthalts

Weitere Zielgrößen:

- unerwünschte Ereignisse der Tests (zum Beispiel Erbrechen)
- gesundheitsbezogene Lebensqualität
- unerwünschte Konsequenzen der Diagnose, zum Beispiel als Folge falsch-positiver oder falsch-negativer Testergebnisse, Überdiagnose oder Labeling
- unerwünschte Folgen einer Intervention mit blutzuckersenkender Intention oder anderer Interventionen, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet sind

4.1.4 Surrogatparameter

Surrogatparameter wurden in diesem Arbeitspapier wie im Abschlussbericht S07-01 gehandhabt. Es wurden, soweit vorhanden, Angaben zu unten stehenden Surrogatparametern extrahiert.

Geburtsgewicht und verwandte Maße

- Geburtsgewicht
- Makrosomie
- LGA („large for gestational age“; schwerer als die 90. Perzentile des Gestationsalters)
- SGA („small for gestational age“; leichter als die 10. Perzentile des Gestationsalters)

Mütterliche Ergebnisse:

- schwangerschaftsbedingte Hypertonie
- spätere Diagnose eines Diabetes mellitus

Kindliche Ergebnisse:

- Frühgeburtlichkeit
- spätere Entwicklung von Übergewicht
- Apgar-Wert

4.1.5 Studientypen

Teilziel 1 „Screening“:

Für Teilziel 1 galten im vorliegenden Arbeitspapier dieselben Vorgaben für Studientypen wie im Abschlussbericht S07-01. Randomisierte kontrollierte Studien (RCTs) liefern für die Bewertung des Nutzens einer Screeningmaßnahme die zuverlässigsten Ergebnisse, da sie, sofern methodisch adäquat und der Fragestellung angemessen durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet sind. Für dieses Arbeitspapier konnten neben randomisierten kontrollierten Studien jedoch auch eindeutig prospektiv geplante nicht randomisierte kontrollierte (Interventions-)Studien herangezogen werden, sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt worden wäre und zwischen den Kollektiven vergleichbare Nebenbedingungen vorgelegen hätten. Weiterhin mussten ausreichend statistische Daten zu Basischarakteristika beider Gruppen verfügbar sein, um die Übertragbarkeit der Ergebnisse auf andere Populationen und den Einfluss wichtiger Störgrößen und Verzerrungen abschätzen zu können. Als prospektiv geplant galt eine Studie dann, wenn vor der Zuteilung der Intervention eine eindeutige Methode für die Zuteilung der Intervention vorgegeben und in einer Methodik (oder einem Studienprotokoll) klar beschrieben war.

Es wurden im vorliegenden Arbeitspapier keine Studien zu Teilziel 1 identifiziert, die diese Bedingungen erfüllten.

Teilziel 2 „Therapie“:

Für Teilziel 2 wurden im vorliegenden Arbeitspapier nur RCTs herangezogen, da nur RCTs die Ergebnissicherheit aufweisen, die notwendig ist, um die Ergebnisse des Abschlussberichts S07-01 modifizieren zu können. In den eingeschlossenen RCTs mussten ausreichend statistische Daten zu Baselinecharakteristika beider Gruppen verfügbar sein, um die Übertragbarkeit der Ergebnisse auf andere Populationen und den Einfluss wichtiger Störgrößen und Verzerrungen abschätzen zu können.

4.1.6 Ein- / Ausschlusskriterien

In die Nutzenbewertung des vorliegenden Arbeitspapiers wurden alle Studien einbezogen, die

- alle nachfolgenden Einschlusskriterien und
- keines der nachfolgenden Ausschlusskriterien erfüllten.

Die Nachrecherche war auf den Zeitraum Oktober 2008 bis Oktober 2009 beschränkt.

Die Einschlusskriterien zu Teilziel 1 sind mit denen des Abschlussberichts S07-01 identisch.

Tabelle 1: Einschlusskriterien Teilziel 1 (S = Screening)

Einschlusskriterien	
S1	unselektierte Schwangere
S2	Screening auf Gestationsdiabetes wie unter 2.1 definiert
S3	Vergleichsgruppe ohne Screening oder mit anderer Screeningstrategie wie unter 2.1 definiert
S4	Studientypen wie unter 4.1.5 definiert
S5	Zielgrößen wie unter 4.1.3 definiert

Die Einschlusskriterien zu Teilziel 2 unterscheiden sich von denen des Abschlussberichts S07-01 dadurch, dass T2 nur Vergleiche Therapie vs. Routine zulässt, während im Abschlussbericht auch Vergleiche von Therapien unterschiedlicher Intensität zugelassen waren. Zudem schränkt T4 nur auf RCTs ein.

Tabelle 2: Einschlusskriterien Teilziel 2 (T = Therapie)

Einschlusskriterien	
T1	Schwangere, bei denen in einem Screeningsetting ein Gestationsdiabetes (Einschränkung der Glukosetoleranz) diagnostiziert wurde
T2	Intervention mit blutzuckersenkender Intention oder Interventionen, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet sind
T3	Vergleichsgruppe ohne eine der in T2 aufgeführten Interventionen
T4	Randomisiert-kontrollierte Studien
T5	Zielgrößen wie unter 4.1.3 definiert

Die nachfolgenden Ausschlusskriterien sind identisch mit denen des Abschlussberichts S07-01.

Tabelle 3: Ausschlusskriterien (alle Teilziele betreffend)

Ausschlusskriterien	
A1	Mehrfachpublikationen ohne relevante Zusatzinformation
A2	Keine Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT-Statements [7] oder im Falle nicht randomisierter Studien vergleichbaren Standards (z. B. TREND [8], STARD [9], STROBE [10,11]) genügt und eine Bewertung der Studie ermöglicht.	

4.2 Informationsbeschaffung

4.2.1 Literaturrecherche

Zur Entwicklung einer angepassten Suchstrategie für dieses Arbeitspapier wurden die in Studienpool A und B des Abschlussberichtes S07-01 eingeschlossenen Studien als Testsatz genutzt. Da alle Studien des Testsatzes durch eine Recherche in den Datenbanken CENTRAL, EMBASE und MEDLINE aufgefunden werden konnten, wurde die Recherche für dieses Arbeitspapier auf diese Datenbanken eingeschränkt.

Die Suche nach relevanten Studien fand in folgenden Quellen statt:

Tabelle 4: Quellen der Information

Quelle	Kommentar
Bibliografische Datenbanken	Suche in folgenden Datenbanken: <ul style="list-style-type: none">▪ CENTRAL▪ EMBASE▪ MEDLINE
Kontaktaufnahmen	<ul style="list-style-type: none">▪ Sofern notwendig mit Autoren, zum Beispiel zur Frage nach nicht publizierten Teilaspekten

Die Suchstrategien sind in Anhang A dokumentiert. Die Suche erfolgte am 26. Oktober 2009.

4.2.2 Identifizierung relevanter Studien

Titel- und Abstractscreening der Ergebnisse der Recherche in bibliografischen Datenbanken

Die durch die Suche in bibliografischen Datenbanken identifizierten Zitate wurden anhand ihres Titels und, sofern vorhanden, Abstracts von 2 Reviewern unabhängig voneinander hinsichtlich ihrer Relevanz bewertet. Publikationen, die beide Reviewer als potenziell relevant

erachteten, wurden anhand ihres Volltextes auf Relevanz geprüft. Zitate, die nur ein Reviewer als potenziell relevant eingestuft hatte, wurden von beiden Reviewern nochmals gesichtet und anschließend nach Diskussion entweder als irrelevant bezeichnet oder ebenfalls anhand ihres Volltextes auf Relevanz geprüft.

Überprüfung potenziell relevanter Volltexte

Die Überprüfung auf Relevanz anhand des Volltextes erfolgte wiederum von 2 Reviewern unabhängig voneinander. Als relevant wurden nach diesem Schritt folgende Studien bezeichnet:

- Studien, die von beiden Reviewern als relevant erachtet wurden
- Studien, die zunächst nur von einem der beiden Reviewer, aber nach anschließender Diskussion von beiden Reviewern als relevant erachtet wurden

4.2.3 Suche nach zusätzlichen Informationen zu relevanten Studien

Es wurden Autoren von Studien kontaktiert, um Fragen zur Bewertung einer Studie zu klären, die anhand der vorliegenden Daten nicht geklärt werden konnten.

4.3 Informationsbewertung

Die Bewertung der eingeschlossenen Studien erfolgte anhand der zur Verfügung stehenden Informationen und hing damit stark von der Qualität der jeweiligen Publikation und weiterer Informationsquellen ab.

Die Bewertung erfolgte in 3 Schritten:

- Extraktion der Daten
- Bewertung der Datenkonsistenz innerhalb der Publikation und zwischen den Informationsquellen
- Bewertung des Verzerrungspotenzials

Datenextraktion

Die Extraktion der Daten publizierter Studien wurde anhand von standardisierten Datenextraktionsbögen vorgenommen. Ein Reviewer führte die Datenextraktion unter Verwendung des Extraktionsbogens durch. Ein zweiter Reviewer überprüfte die Extraktion. Diskrepanzen in der Bewertung wurden durch Diskussion zwischen den Reviewern aufgelöst.

Bewertung des Verzerrungspotenzials der Ergebnisse auf Studien- und Endpunktebene

Um darzustellen, mit welcher Sicherheit sich ein Effekt oder das Fehlen eines Effekts aus einer Studie ableiten lässt, erfolgte für die in die Nutzenbewertung eingeschlossenen Studien eine Bewertung des Verzerrungspotenzials der Ergebnisse.

Diese erfolgte sowohl für die Studie insgesamt als auch für die einzelnen Endpunkte. Für die Bewertung des Verzerrungspotenzials der Ergebnisse auf Studienebene wurden insbesondere die folgenden Angaben zum Studiendesign und zur Studiendurchführung systematisch extrahiert:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuweisung
- Patientenfluss
- Methodik zur Verblindung der Patientinnen sowie der Behandler

Falls es weitere Aspekte gab, die Bias verursachen können (zum Beispiel Dateninkonsistenzen), wurden diese ebenfalls einbezogen.

Für die Bewertung des Verzerrungspotenzials der Ergebnisse der Endpunkte wurden insbesondere folgende Angaben systematisch extrahiert:

- Methodik zur Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung, die nur das Verzerrungspotenzial für diesen Endpunkt beeinflusst

Falls es weitere Aspekte gab, die Bias verursachen können, wurden diese ebenfalls einbezogen. Anhand der Bewertungen der oben genannten Aspekte wurde das Verzerrungspotenzial der Ergebnisse der Studien in die Ausprägungen „hoch“ und „niedrig“ eingeteilt.

Die Bewertungen der Einzelaspekte und die Gesamteinschätzung des Verzerrungspotenzials auf Studienebene wurden in Tabellen zusammenfassend dargestellt. Die Bewertung des Verzerrungspotenzials einzelner Endpunkte wurde nur dann berichtet, wenn sie von der Einschätzung der Gesamtstudie abwich.

Die Ausprägungen wurden vorab wie folgt definiert: Ein „niedriges Verzerrungspotenzial“ lag vor, wenn davon ausgegangen wurde, dass das Ergebnis (des Endpunktes bzw. der Studie insgesamt) durch die möglicherweise aufgezeigten Defizite im Studiendesign bzw. in der Studiendurchführung nicht wesentlich verzerrt / beeinflusst wird. Ein „hohes Verzerrungs-

potenzial“ lag vor, wenn davon ausgegangen wurde, dass das Ergebnis (des Endpunktes bzw. der Studie) insgesamt infrage gestellt wird.

4.4 Informationssynthese und -analyse

4.4.1 Charakterisierung der Studien

Als zentrale Studiencharakteristika wurden zunächst das Design der jeweiligen Studie und deren Eignung für den Bewertungsprozess extrahiert. Außerdem wurden Charakteristika wie Stichprobe, Setting, Art der Behandlung, Ausschluss- und Einschlussgründe, die Behandlung selbst und die Zielparmeter dargestellt. Die Populationen wurden durch die Basischarakteristika Alter, Anteil der Erstgebärenden, Schwangerschaftswoche, Gewicht / BMI, Ethnie und Nüchtern-BG bei Einschluss charakterisiert. Die Studienqualität wurde gesondert beschrieben.

Für die eingeschlossenen Studien wurden die Zielgrößen extrahiert wie in 4.1.3 beschrieben.

4.4.2 Meta-Analyse

Sofern die Studienlage es inhaltlich und methodisch als sinnvoll erscheinen ließ, erfolgten quantitative Zusammenfassungen der Einzelergebnisse mithilfe von Meta-Analysen gemäß den Methoden des Instituts. Für die statistische Auswertung wurden – sofern möglich – die Ergebnisse aus Intention-to-Treat-Analysen verwendet, so wie sie in den Publikationen beschrieben sind.

Die Meta-Analysen erfolgten zunächst auf der Basis von Modellen mit zufälligen Effekten nach der Methode von DerSimonian und Laird. Als Effektmaß wurde das Odds Ratio verwendet. Trat in einer der Behandlungsgruppen das Ereignis bei keinem Patienten ein, so wurde in der zugrunde liegenden Vierfeldertafel zu jeder Zellohäufigkeit ein Korrekturwert von 0,5 addiert. Bei sehr geringen Ereignisraten (< 1 %) wurden die Meta-Analysen auf der Basis des Peto Odds Ratios (in einem Modell mit festen Effekten) durchgeführt.

Da häufig die Anzahl der zur Verfügung stehenden Studien gering war, wurden zum Vergleich auch Meta-Analysen auf Basis von Modellen mit festen Effekten nach der Mantel-Haenszel-Methode berechnet. Die Berechnungen mit festen Effekten haben jedoch in keinem Fall abweichende Ergebnisse ergeben. Deshalb sind in diesem Arbeitspapier nur die Meta-Analysen auf Basis von Modellen mit zufälligen Effekten oder auf der Basis des Peto Odds Ratios wiedergegeben.

Für Heterogenitätsuntersuchungen wurden Heterogenitätstests durchgeführt, das I^2 -Maß berechnet sowie die Lage der Effektschätzer in den Forest-Plots beurteilt. Falls der Heterogenitätstest ein Ergebnis von $p < 0,2$ lieferte, wurde kein gepoolter Effektschätzer dargestellt.

4.4.3 Sensitivitätsanalyse

Sensitivitätsanalysen fanden in diesem Arbeitspapier nicht statt (siehe Abschnitt 4.3).

4.4.4 Subgruppenanalyse

Da in der eingeschlossenen Studie keine Ergebnisse für vergleichbare Subgruppen berichtet wurden und sich keine Angaben ableiten ließen, die eine Bildung von Subgruppen erlaubt hätten, fanden diese Analysen nicht statt.

5 Ergebnisse

Ein Screening auf Gestationsdiabetes kann positive und negative Auswirkungen haben, auch auf Schwangere ohne Gestationsdiabetes. Teilziel 1 (Abschnitt 5.1) dieses Arbeitspapiers war deshalb die Recherche und Bewertung von Studien, die Auswirkungen eines Screeningprogramms auf alle Teilnehmerinnen betrachten. Dieses Teilziel umfasst auch die vergleichende Bewertung der Effektivität unterschiedlicher Screeningprogramme. Im Idealfall würde eine klare Evidenz zu diesem Teilziel schon für eine Nutzenbewertung ausreichen.

Alternativ zu Teilziel 1 „Screening“ wurde in Abschnitt 5.2 versucht, Hinweise auf einen Nutzen aus Analysen von Studien abzuleiten, die in Teilziel 2 „Therapie“ eingeschlossen waren. Dieses Teilziel beinhaltete die Bewertung der Effektivität von Therapien unterschiedlicher Intensität für Schwangere mit GDM anhand vergleichender Studien.

5.1 Teilziel 1 „Screening“

Ziel dieses Berichtsteils ist die Recherche und Bewertung von vergleichenden Studien, die Auswirkungen eines Screeningprogramms auf alle Teilnehmerinnen betrachten. Dieses Teilziel umfasst auch die vergleichende Bewertung der Effektivität unterschiedlicher Screeningprogramme.

5.1.1 Ergebnis der Literaturrecherche

Abbildung 1 zeigt die Ergebnisse der gemeinsamen systematischen Literaturrecherche für die Teilziele 1 und 2 in bibliografischen Datenbanken und die des Literaturscreenings gemäß den Ein- / Ausschlusskriterien.

Die Recherchen ergaben nach der Dublettenbereinigung insgesamt 260 Treffer. Für Teilziel 1 konnte kein potenziell relevanter Treffer identifiziert werden.

Die verwendeten Suchstrategien sind in Anhang A dargestellt.

5.2 Teilziel 2 „Therapie“

Ziel dieses Teilziels ist die Recherche und Bewertung von vergleichenden Studien, die die Effektivität von Therapien für Schwangere mit GDM betrachten.

5.2.1 Ergebnis der Literaturrecherche

Abbildung 1 zeigt die Ergebnisse der systematischen Literaturrecherche in bibliografischen Datenbanken und des Literaturscreenings gemäß den Ein- / Ausschlusskriterien.

Die Recherchen ergaben nach der Dublettenbereinigung insgesamt 260 Treffer. Von den Treffern wurden im ersten Screening 256 als nicht relevant ausgeschlossen. Vier Publikationen wurden im Volltext überprüft. Im zweiten Screening wurden davon 3 Publikationen ausgeschlossen. Die Zitate der im Volltext gesichteten, aber nicht relevanten Publikationen finden sich mit Angabe des Ausschlussgrunds in Anhang B.

Die verwendeten Suchstrategien sind in Anhang A dargestellt.

Letztlich verblieb 1 Studie, die nach der Aktualisierungsrecherche in den Studienpool A aufgenommen wurde. Zu dieser Studie konnten 2 weitere, ältere Publikationen identifiziert werden, in denen das Design vorgestellt wurde (siehe Tabelle 5).

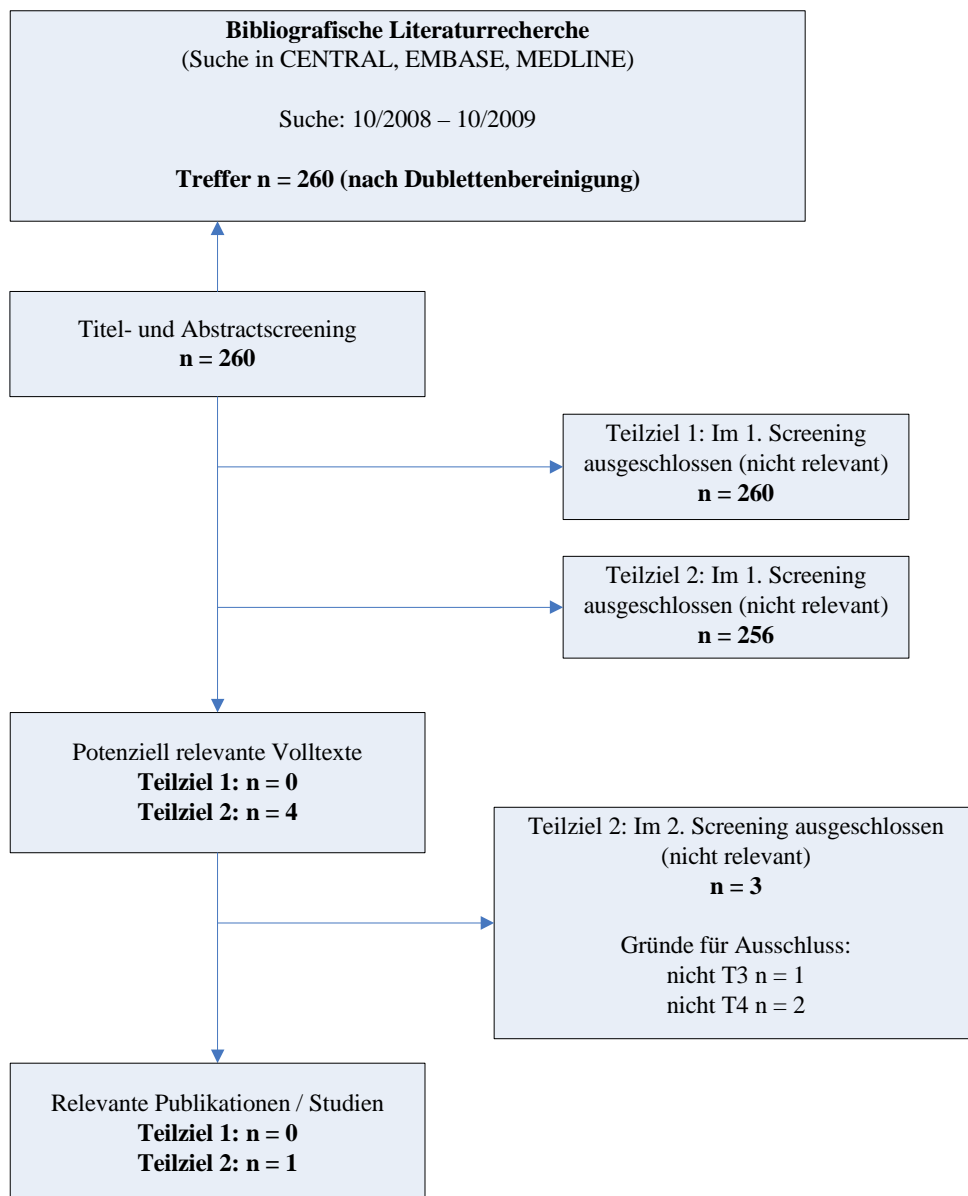


Abbildung 1: Teilziele 1 und 2 – Ergebnisse der Aktualisierungsrecherche

5.2.1.1 Resultierender Studienpool

Die Aktualisierungsrecherche hat außer Landon et al [6] keine weitere Studie identifiziert. Tabelle 5 enthält den aktualisierten Studienpool A.

Tabelle 5: Studienpool A – Ergebnisse der Literaturbeschaffung

Studie	Vollpublikation/en	einzuschließen in Arbeitspapier
Landon 2009	Landon et al. N Engl J Med 2009;361(14):1339-48 [6] Landon et al. Diabetes Care 2007; 30 (Suppl 2): S194-9 [4] Landon et al. J Matern Fetal Neonatal Med. 2002; 11(4): 226-31 [5]	ja / neu
Bonomo 2005	Bonomo et al. Diabet Med 2005; 22(11): 1536-1541 [12]	ja / bereits im Abschlussbericht enthalten
Crowther 2005	Crowther et al. N Engl J Med 2005; 352(24): 2477-2486 [13] Athukorala et al. Aust N Z J Obstet Gynaecol 2007; 47(1): 37-41 [14] Moss et al. BMC Pregnancy Childbirth 2007; 7: 27 [15]	ja / bereits im Abschlussbericht enthalten
Langer 1989	Langer et al. Am J Obstet Gynecol 1989; 161(3): 646-653 [16] Herbison P et al. Am J Obstet Gynecol 1995; 172(5): 1642-1643 [17]	ja / bereits im Abschlussbericht enthalten
O'Sullivan 1966	O'Sullivan et al. Obstetrics & Gynecology 1966; 27: 683-689 [18] O'Sullivan ; Buchbeitrag Churchill Livingstone; 1975. S. 195-204 [19] O'Sullivan et al. Diabetes Care 1980; 3(3): 482-485 [20]	ja / bereits im Abschlussbericht enthalten

5.2.1.2 Studienpool A: Charakteristika der in die Bewertung eingeflossenen randomisierten Studien

5.2.1.2.1 Studiendesign und Studienpopulationen

In Tabelle 6 und Tabelle 7 sind zu dem aktualisierten Studienpool A Angaben zum Design der randomisierten Studien und zu den Basisdaten der untersuchten Populationen zusammengefasst.

Studiendesign: Landon 2009, Bonomo 2005, Crowther 2005, Langer 1989 und O'Sullivan 1966 waren randomisiert-kontrollierte Interventionsstudien mit parallelen Gruppen. In

Landon 2009 und Crowther 2005 waren die Frauen in der Kontrollgruppe und deren Behandler für die Ergebnisse des oGTT verblindet, die übrigen Studien wurden in einem ausschließlich offenen Design durchgeführt.

Population / Setting: Die Studienzentren lagen in Europa, in Australien und den USA; 2 RCTs waren monozentrisch und 3 multizentrisch. Alle Studien fanden in der Ambulanz von Geburtskliniken statt. Die älteste Studie wurde 1966 publiziert, die jüngste 2009. Die Zahl der Teilnehmerinnen reichte von 126 bis 1000 Frauen.

Basischarakteristika: Das mittlere Alter der Frauen war in allen Studien angegeben und lag zwischen 28 und 31 Jahren. Der Anteil der Erstgebärenden lag in den 5 RCTs zwischen 11 % und 49 %. Drei der 5 Studien berichteten über die Dauer der Schwangerschaft bei Studieneinschluss und in 4 von 5 Studien fanden sich Angaben zum BMI. In 2 Studien waren die Frauen mit einem mittleren BMI von 23 bzw. 26 relativ schlank, in der dritten lag der BMI bei etwa 30 kg/m². Angaben zur Ethnie fanden sich in 4 der 5 RCTs. Der Anteil der Kaukasierinnen lag zwischen etwa 25 % und etwa 100 % (siehe Tabelle 7).

Ein- / Ausschlusskriterien: Das wesentliche Einschlusskriterium war in allen Studien die Diagnose einer gestörten Glukosetoleranz mittels oGTT. Allerdings unterschieden sich die Studien darin, wie eine Vorauswahl aufgrund von Risikofaktoren oder eines GCT stattfand und welche Glukosegrenzwerte für den Einschluss festgelegt wurden. Bonomo 2005 schloss Frauen ein, die lediglich einen auffälligen Einstunden-GCT-Wert über 140 mg/dl aufwiesen, bei denen der nachfolgende 100-g-oGTT nach Carpenter-Coustan- / ADA-Kriterien jedoch unauffällig war. Diese Frauen wiesen also eine milde Störung der Glukosetoleranz auf. Langer 1989 untersuchte Frauen, die sowohl einen positiven Einstunden-GCT als auch genau einen auffälligen Wert im Dreistunden-100-g-oGTT nach NDDG-Kriterien hatten, sodass die Störung der Glukosetoleranz bereits ausgeprägter war. Auch in O'Sullivan 1966 wurden die Frauen durch einen Einstunden-GCT, gefolgt von einem Dreistunden-100-g-oGTT ausgewählt. Die Grenzwerte waren selbstdefiniert. Crowther 2005 verwendete zwar die WHO-Kriterien für einen 75-g-oGTT, aber im Unterschied zur WHO-Empfehlung fand eine Vorselektion der Frauen durch einen Einstunden-50-g-GCT (93 % der in Crowther 2005 eingeschlossenen Frauen) oder die Abfrage von Risikofaktoren (7 % der in Crowther 2005 eingeschlossenen Frauen) statt. Insgesamt ist davon auszugehen, dass in Crowther 2005 vor allem Schwangere mit einer ausgeprägten Glukosetoleranzstörung bis an die Grenze zum manifesten Diabetes mellitus eingeschlossen wurden. In Landon 2009 wurden Schwangere eingeschlossen, die durch einen Einstunden-GCT vorselektioniert worden waren und 2 auffällige Werte in einem Dreistunden-100-g-oGTT nach Carpenter-Coustan- / ADA-Kriterien aufwiesen.

Allen RCTs in Pool A ist damit gemeinsam, dass bei der Selektion der Teilnehmerinnen ein GCT als Vortest eingesetzt worden war; selbst in Crowther 2005 wurden so 93 % der Frauen eingeschlossen. In Bonomo 2005 und Crowther 2005 lag der Grenzwert für einen positiven

GCT bei ≥ 140 mg/dl, in Langer 1989 und O'Sullivan 1966 lag er bei > 130 mg/dl, in Landon 2009 bei ≥ 135 mg/dl.

Detaillierte Angaben zu den Diagnosekriterien und Grenzwerten enthält Tabelle 8.

Therapien: Allen RCTs des Studienpools A ist gemeinsam, dass in den Kontrollgruppen eine „Routinebehandlung“ stattfand: Für die Schwangeren war also zum Beispiel keine Gestationsdiabetes-spezifische Diätberatung oder keine besondere geburtshilfliche Betreuung vorgesehen. In Landon 2009 erhielten alle Frauen die Information, welcher Gruppe (IG oder KG) sie zugeteilt wurden, die Ergebnisse des oGTT blieben jedoch verblindet. In Bonomo 2005 wurden die Patientinnen in der Kontrollgruppe hinsichtlich der Diagnose einer Glukosetoleranzstörung beschwichtigt und erhielten keine spezielle Betreuung, Diät oder pharmakologische Behandlung. In Crowther 2005 wurde den Frauen in der Kontrollgruppe die Information gegeben, dass sie keinen GDM aufwiesen, obwohl sie die WHO-Kriterien für eine Störung der Glukosetoleranz (und späteren GDM) erfüllten. Diese Frauen hatten zu diesem Zeitpunkt bereits eine erste Screeningstufe absolviert. Falls die Glukosetoleranzstörung jedoch als Hinweis auf einen möglichen Diabetes mellitus gesehen wurde, durfte der Verdacht entsprechend dem Standardvorgehen im jeweiligen Zentrum abgeklärt und bei Bestätigung behandelt werden. In Langer 1989 wurde den Frauen der Kontrollgruppe angeraten, ihre Essgewohnheiten nicht zu ändern. In O'Sullivan 1966 fand in der Kontrollgruppe lediglich eine pränatale Routinebetreuung statt. Da die Teilnehmerinnen der Kontrollgruppen bis auf die in Crowther 2005 und Landon 2009 nicht verblindet waren, kann nicht ausgeschlossen werden, dass das Wissen um die Blutzuckerwerte Verhaltens- und Betreuungsänderungen induziert hat. In den Interventionsgruppen wurde in allen Studien als erste wesentliche Maßnahme eine Diät eingesetzt. Die Studien unterscheiden sich durch den zusätzlichen Einsatz von Insulin: In Bonomo 2005 verwendeten die Frauen kein Insulin. In Crowther 2005, Langer 1989 und Landon 2009 begannen sie bei unzureichender BG-Einstellung mit Insulin. In O'Sullivan 1966 setzten alle Frauen der Interventionsgruppe zusätzlich Insulin ein. Für weitere Informationen über die Art der Behandlung siehe Tabelle 8.

Zielgrößen: In der Auswahl der Zielgrößen sind die Studien sehr heterogen. Das ist zum Teil historisch bedingt. O'Sullivan 1966 untersuchte als mütterliche Zielgröße die spätere Entwicklung eines Diabetes mellitus, während jüngere Studien vor allem Geburtsergebnisse wie den Anteil der Kaiserschnitte und die Art der Geburt erfassten. Der primäre Endpunkt wurde von Bonomo 2005, Crowther 2005 und Landon 2009 dezidiert genannt, in Langer 1989 und O'Sullivan 1966 fanden sich keine Angaben zu primären Zielkriterien (siehe Tabelle 6). Zudem berichteten alle Studien über die kindliche Mortalität, hinzu kamen Geburtskomplikationen wie Schulterdystokie und Stoffwechselstörungen wie Hypoglykämie. Außerdem berichteten alle Studien über körpermaßassozierte Zielgrößen, entweder als durchschnittliches Geburtsgewicht oder als Anteil der Kinder mit Makrosomie oder LGA.

Tabelle 6: Studienpool A – Übersicht über die eingeschlossenen randomisierten Studien

Studie	Studiendesign	Vergleich	Zahl der Schwangeren	Land / Versorgungskontext	Relevante Zielgrößen	
					Mutter	Kind
Landon 2009	RCT teilweise verblindet multizentrisch	Diät / BG-Selbstmessung / Insulin vs. Routinebehandlung	IG: 485 KG: 473	USA Krankenhaus	Gestationsalter bei Geburt, Gewichtszunahme von Studieneinschluss bis Geburt, Bluthochdruck in der Schwangerschaft, Präeklampsie, Schulterdystokie, Kaiserschnitt, operative Entbindung, Dauer des Klinikaufenthalts	primärer Endpunkt: Fehlgeburt / neonatale Mortalität, neonatale Hypoglykämie, Hyperbilirubinämie, neonatale Hyperinsulinämie (C-Peptid > 95. Perzentile) sekundäre Endpunkte: Komponenten des primären Endpunkts, Frühgeburten, Geburtstraumata, Geburtsgewicht, Makrosomie, LGA, SGA, klinische Hypoglykämie, Hyperbilirubinämie mit Fototherapie, Atemnotsyndrom, Aufnahme auf Intensivstation
Bonomo 2005	RCT offen monozentrisch	Diät vs. Routine	IG: 150 KG: 150	Italien Krankenhaus	Gestationsalter Kaiserschnitt	Geburtsgewicht, LGA, Makrosomie, SGA Hypoglykämie, Hyperbilirubinämie, Polyzythämie Aufnahme auf eine Intensivstation Apgar-Wert
Crowther 2005	RCT teilweise verblindet multizentrisch	Diät / BG-Selbstmessung / Insulin vs. Routine	IG: 490 KG: 510	Australien ^a Krankenhaus	primäre Endpunkte: Einleitung der Geburt, Kaiserschnitt, Lebensqualität sekundäre Endpunkte: Anzahl pränataler Kontakte, Art der Geburt, Anzahl der Krankenhausaufenthalte, Hypertonie, Gestationsalter	primäre Endpunkte: kombinierter Endpunkt (Tod, Schulterdystokie, Knochenfraktur, Nervenlähmung), Aufnahme auf neonatale Pflegestation, Hyperbilirubinämie sekundäre Endpunkte: Komponenten des primären kombinierten Endpunkts, Geburtsgewicht

(Fortsetzung)

Tabelle 6: Studienpool A – Übersicht über die eingeschlossenen randomisierten Studien (Fortsetzung)

Studie	Studiendesign	Vergleich	Zahl der Schwangeren	Land / Versorgungskontext	Relevante Zielgrößen	
					Mutter	Kind
Langer 1989	RCT offen monozentrisch	Diät / Insulin vs. Routine	IG: 63 KG: 63	USA Krankenhaus	Kaiserschnitt	LGA, SGA Frühgeburtlichkeit Hypoglykämie, Hyperbilirubinämie, Hypokalzämie, Polyzythämie, Atemnotsyndrom Aufnahme auf eine Intensivstation
O'Sullivan 1966	RCT offen multizentrisch (n = 2)	Diät / Insulin vs. Routine	IG: 307 KG: 308	USA Krankenhaus	spätere Diagnose eines Diabetes mellitus	Mortalität Makrosomie kongenitale Anomalien Geburtsgewicht Frühgeburtlichkeit
RCT: randomisiert-kontrollierte Studie; IG: Interventionsgruppe; KG: Kontrollgruppe; BG: Blutglukosekonzentration a: 46 von 1000 Frauen wurden in 3 Kliniken in Großbritannien rekrutiert						

Tabelle 7: Studienpool A – Basischarakteristika der Studienpopulationen

Studie / Gruppen	N	Alter Jahre (SD)	Erstgebärende (%)	SSW ^a	Gewicht kg (SD) ^a	BMI kg/m ² (SD) ^a	Nüchtern-BG mg/dl (SD)	Ethnie
Landon 2009					k. A.			11,5 % / 11,4 % „Schwarz“ 25,4 % / 25,2 % „Weiß“ 4,5 % / 5,9 % „Asiatisch“ 57,9 % / 56 % „Hispanic“ 0,6 % / 1,5 % „Andere“
	IG: 485	29,2 (5,7)	21	28,8 (1,6)		30,1(5,0)	86,6 (5,7)	
	KG: 473	28,9 (5,6)	26	28,9 (1,5)		30,2 (5,1)	86,3 (5,7)	
Bonomo 2005				k. A.	k. A.			100 % Kaukasierinnen
	IG: 150	31 (5)	45			23 (4) ^b	85,9	
	KG: 150	31 (5)	42			23 (5) ^b	84,3	
Crowther 2005								73 % / 78 % Kaukasierinnen 19 % / 14 % Asiatinnen 9 % / 8 % andere
	IG: 490	31 (5)	43	29 (median) (IQR:28–30)	k. A.	27 (median) (IQR:23–31) ^c	86,5 (12,6)	
	KG: 510	30 (6)	49	29 (median) (IQR:28–30)		26 (median) (IQR:23–31) ^c	86,5 (10,8)	
Langer 1989					k. A.	k. A. ^d		36 % / 33 % Kaukasierinnen 33 % / 33 % Lateinamerikanerinnen 30 % / 33 % Afroamerikanerinnen
	IG: 63	31 (5)	29	31 (3)			93	
	KG: 63	28 (6)	32	31 (3)			90	
O’Sullivan 1966				k. A.	k. A. ^f	k. A.	k. A.	k. A.
	IG: 307	30 (k. A.)	14 ^e					
	KG: 308	31 (k. A.)	11 ^e					
IG: Interventionsgruppe; KG: Kontrollgruppe; IQR: Interquartilsabstand; k. A.: keine Angabe; BMI: Body-Mass-Index; BG: Blutglukosekonzentration; SSW: Schwangerschaftswoche bei Einschluss a: Angabe als Mittelwert und Standardabweichung (SD), sofern nicht anders angegeben, gerundet b: Zeitpunkt: Studieneintritt bzw. GDM-Diagnose c: erstes Trimester d: IG: 38 % hatten BMI $\geq 27\text{kg/m}^2$; KG: 41 % hatten BMI $\geq 27\text{kg/m}^2$ e: errechnet f: laut Publikation kein Unterschied zwischen den Gruppen								

Tabelle 8: Studienpool A – Beschreibung der Interventionen und der Ein- und Ausschlusskriterien

Studie	Behandlung	Strategie zur Diagnose von GDM	Grenzwerte für Einschluss	wesentliche Ausschlusskriterien
Landon 2009	IG: Diät / tgl. BG-Selbstmessung / Insulin, wenn NüBG \geq 95 mg/dl oder 2 hpostprandial \geq 120 mg/dl KG: Routinebehandlung	1. 50-g-GCT 24.–31. SSW 2. Wenn 50-g-GCT-1-h-Wert 135–200 mg/dl, dann 3-h-100-g-oGTT	pos. 1-h-50-g-GCT und pos. 3-h-100-g-oGTT Kriterien: NüBG < 95 mg/dl und \geq 2 der folgenden Werte überschritten: 1 h \geq 180 mg/dl 2 h \geq 155 mg/dl 3 h \geq 140 mg/dl	- vorbestehender Diabetes mellitus - auffälliger GCT vor der 24. SSW - GDM in früherer Schwangerschaft - frühere Fehlgeburt - frühere fetale Fehlbildungen - aktuelle Mehrlingschwangerschaft - Risiko für vorzeitige Entbindung - Asthma - chronischer Bluthochdruck - Einnahme von Kortikosteroiden
Bonomo 2005	IG: Diät (24–30 kcal/kg je nach Körpergewicht) plus ambulantes Management mit 14tägigen Visiten; BG-Zielwerte: Nü < 90 mg/dl und 2 h postprandial < 121 mg/dl KG: Routinebehandlung	1. Erhebung von RF 2a. RF-pos. Frauen: 50-g-GCT bei Erstvorstellung 2b. RF-neg. Frauen: 50-g-GCT in der 24.–28. SSW 2c. RF-pos. Frauen mit neg. GCT: erneuter 50-g-GCT in der 30.–34. SSW 3. Wenn 50-g-GCT-1-h-Wert \geq 140 mg/dl, dann 100-g-oGTT	pos. 1-h-50-g-GCT aber neg. 100-g-oGTT nach Carpenter-Coustan-Kriterien	- normaler GCT - ein abnormaler Wert im oGTT - diagnostizierter GDM gemäß Carpenter-Coustan-Kriterien

(Fortsetzung)

Tabelle 8: Studienpool A – Beschreibung der Interventionen und der Ein- und Ausschlusskriterien

Studie	Behandlung	Strategie zur Diagnose von GDM	Grenzwerte für Einschluss:	wesentliche Ausschlusskriterien
Crowther 2005	<p>IG: a: individuelle professionelle Diätberatung b: 4x tgl. BG-Selbstmessung bis BG über 2 Wochen im Zielbereich: NüBG: 63–99 mg/dl, präprandial \leq 99 mg/dl, 2 h postprandial \leq 126 mg/dl; danach tgl. BG-Selbstmessung c: Insulintherapie bzw. -anpassung erfolgte, falls - ein Messwert \geq 162 mg/dl lag - bis zur 35. SSW: falls in 2 Wochen 2 Messwerte NüBG \geq 99 mg/dl oder postprandial \geq 126 mg/dl lagen - nach der 35. SSW: falls in 2 Wochen 2 Messwerte NüBG \geq 99 mg/dl oder postprandial \geq 144 mg/dl waren KG: Standardbetreuung; bei Hinweis auf Diabetes weitere Abklärung erlaubt und Therapie möglich.</p>	<p>1. Erhebung von RF oder 50-g-GCT in der 16.–30. SSW 2. RF-pos. Frauen oder 50-g-GCT-1-h-Wert \geq 140 mg/dl; 75-g-oGTT</p>	<p>75-g oGTT in der 24. bis 34. SSW: ven. Plasma nüchtern: $<$ 140 mg/dl und 2-h-Wert: 140–198 mg/dl (ab 1998: jeder abnormale Glukosewert bis 198 mg/dl) Ein- oder Zwillingsschwangerschaften zwischen 16. und 30. SSW</p>	<p>- BG ven.Plasma: \geq 198 mg/dl - vorangegangener, behandelter Gestationsdiabetes - chronische systemische Erkrankung (außer essenzielle Hypertonie)</p>
Langer 1989	<p>IG: Diät in Abhängigkeit vom BMI vor der Schwangerschaft plus bei Bedarf Insulin (Beginn bei NüBG $>$ 95 mg/dl mit 0,7 U Insulin/ kg KG als MDI-Regime) KG: Essgewohnheiten unverändert fortgesetzt</p>	<p>1. 50-g-GCT 2. wenn 50-g-GCT-1-h-Wert $>$ 130 mg/dl, dann 100-g-oGTT</p>	<p>1 positiver Wert im 3-h-100-g-oGTT nach NDDG 1979[21]</p>	<p>- normaler oGTT - kein oder $>$ 1 erhöhter Wert nach NDDG 1979 [21]</p>
O’Sullivan 1966	<p>IG: Diät (30 kcal/kg Idealgewicht) plus Insulin: 10 U NPH-Insulin 1x tgl. morgens; Dosisanpassung nach Glukosurie plus „printed routine obstetric and dietary instructions“ KG: Routinebehandlung plus „printed routine obstetric and dietary instructions“</p>	<p>1. Erhebung von RF oder 50-g-GCT 2. wenn RF-positiv oder 50-g-GCT-1-h-Wert $>$ 130 mg/dl: dann 100-g-oGTT 3. wenn erster GCT neg., aber RF-pos: dann 2. GCT im nachfolgenden Trimester</p>	<p>positiver GCT oder ein Risikofaktor plus positiver 3-h-100-g-oGTT^a</p>	<p>- vorbestehender Diabetes - BG $>$ 300 mg/dl - Diabetes-Symptome - \geq 37. SSW</p>
<p>IG: Interventionsgruppe; KG: Kontrollgruppe; MDI: multiple daily injections; RF: Risikofaktor; Nü: nüchtern; BG: Blutglukosekonzentration; GCT: Glukose-Challenge-Test; oGTT: oraler Glukosetoleranztest; NüBG: Nüchtern-Blutglukosekonzentration a: \geq 2 Grenzwerte überschritten; venöses Vollblut: NüBG \geq 110 mg/dl, 1 h \geq 170 mg/dl, 2 h \geq 120 mg/dl, 3 h \geq 110 mg/dl</p>				

5.2.1.3 Studienpool A: Verzerrungspotenzial der Studien

Zur Einschätzung des Verzerrungspotenzials wurden die 5 in Studienpool A eingeschlossenen Studien anhand standardisierter Dokumentations- und Extraktionsbögen bewertet. Eine Übersicht über diese Bewertung gibt Tabelle 9. Insgesamt wurde das Verzerrungspotenzial nur bei Landon 2009 und Crowther 2005 als niedrig eingestuft, das der übrigen Studien wurde als hoch bewertet.

Randomisierung und Zuteilung: Nur in Landon 2009 und Crowther 2005 waren die Randomisierung und die Zuteilung so nachvollziehbar beschrieben, dass sie als adäquat beurteilt wurden. Bei 3 Studien blieb die Beschreibung unklar.

Verblindung: In Landon 2009 und Crowther 2005 waren die Frauen in der Kontrollgruppe und ihre Behandler für die Ergebnisse des oGTT verblindet. In Crowther 2005 wurde den Frauen der Kontrollgruppe mitgeteilt, es liege kein Gestationsdiabetes vor. In Landon 2009 wurde den Frauen mitgeteilt, in welcher Gruppe sie sich befinden, die konkreten Ergebnisse des oGTT blieben jedoch verblindet. In diesen beiden Studien wurde den Kontrollgruppen auch ein Überschuss von Frauen ohne GDM zugeteilt, sodass in diesen Gruppen unklar blieb, welche Frau tatsächlich das Einschlusskriterien „Diagnose GDM“ erfüllte. Die Frauen in der Interventionsgruppe und ihre Behandler waren hingegen nicht gegenüber der Diagnose bzw. Gruppenzuteilung verblindet. Die Verblindung nur einer Gruppe führte dazu, dass Effekte eines möglichen Labelings als Teil der Intervention betrachtet wurden. In den übrigen Studien ist keine Verblindung beschrieben. In Landon 2009 ist beschrieben, dass alle Fälle von Bluthochdruck und Schulterdystokie verblindet zentral begutachtet wurden, um die Diagnose zu bestätigen.

Ergebnisunabhängige Berichterstattung: Nur in Landon 2009 fanden sich Hinweise auf eine ergebnisgesteuerte Berichterstattung. In den Designpublikationen [4,5] sind einige Endpunkte genannt (operative delivery, gestational hypertension, length of stay), die in der Ergebnispublikation [6] nicht oder als kombinierter Endpunkt berichtet wurden. Andere in der Ergebnispublikation berichtete Endpunkte (zum Beispiel fat mass, BMI bei Entbindung) waren wiederum in den Designpublikationen nicht erwähnt. Allerdings ist keiner dieser Endpunkte von zentraler Bedeutung für die Interpretation der Studie und es gibt keinen Hinweis darauf, dass auch wichtige Endpunkte ergebnisgesteuert berichtet wurden.

Die Hinweise auf eine ergebnisabhängige Berichterstattung hatten deshalb keinen negativen Einfluss auf die Einschätzung des Verzerrungspotenzials von Landon 2009.

Transparenz des Patientenflusses: Bei Landon 2009, Bonomo 2005, Crowther 2005 und O’Sullivan 1966 ist der Patientenfluss vollständig nachvollziehbar. Bei Langer 1989 fehlen Angaben. Drei Studien (Crowther 2005, Langer 1989, O’Sullivan 1966) sind nach dem Intention-To-Treat-Prinzip ausgewertet. In Landon 2009 fließen nicht alle randomisierten Frauen in die Analyse ein, die Auswirkungen auf das Verzerrungspotenzial sind je nach

Endpunkt jedoch unterschiedlich zu bewerten und stellen die Ergebnisse der Studie nicht generell infrage.

Bewertung des Verzerrungspotenzials: Lediglich für Landon 2009 und Crowther 2005 lassen sich für fast alle geprüften Aspekte angemessene Vorkehrungen zur Minimierung von Bias entnehmen. Ausnahmen sind die fehlende Verblindung der Frauen der Interventionsgruppe und die unklare Verblindung der Endpunkterheber. Insgesamt wird das Verzerrungspotenzial von Landon 2009 und Crowther 2005 auf Studienebene und für die meisten Endpunkte als niedrig / niedrig eingeschätzt. Für die 3 übrigen RCTs lautet die Einstufung hoch / hoch, weil wichtige Vorkehrungen gegen Bias nicht beschrieben waren.

Tabelle 9: Studienpool A – Einschätzung des Verzerrungspotenzials der eingeschlossenen Studien

Studie	studienbezogene Bewertung						endpunktbezogene Bewertung			Studie / Endpunkte
	Randomisierung adäquat	Verdeckung adäquat	Verblindung Frauen	Verblindung Behandler	ergebnisunabhängige Berichterstattung	sonstige Aspekte	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	sonstige Aspekte	Verzerrungspotenzial
Landon 2009	ja	ja	IG: nein KG: ja ^a	IG: nein KG: ja ^a	nein ^b	Patientenfluss transparent	ja für Hypertonie und Schulterdys-tokie unklar für alle anderen Endpunkte	nein ^c	nein	niedrig / niedrig ^d
Bonomo 2005	unklar	unklar	nein	nein	ja	1. Analysen enthalten Vergleiche mit dritter, nicht randomisierter Gruppe 2. Patientenfluss transparent	unklar für alle Endpunkte	nein für alle Endpunkte	nein	hoch / hoch
Crowther 2005	ja	ja	IG: nein KG: ja ^a	IG: nein KG: ja ^a	ja	Patientenfluss transparent	unklar für alle Endpunkte	ja / nein für QoL ^e	nein	niedrig / niedrig ^f
Langer 1989	unklar	unklar	nein	nein	ja	1. Analysen enthalten Vergleiche mit dritter, nicht randomisierter Gruppe 2. Patientenfluss transparent	unklar für alle Endpunkte	ja für alle Endpunkte	VZ für Kaiserschnitt rate: hoch	hoch / hoch
O’Sullivan 1966	unklar	unklar	nein	nein	ja	Patientenfluss nicht transparent	unklar für alle Endpunkte	ja für alle Endpunkte	nein	hoch / hoch

(Fortsetzung)

Tabelle 9: Studienpool A – Einschätzung des Verzerrungspotenzials der eingeschlossenen Studien (Fortsetzung)

IG: Interventionsgruppe; KG: Kontrollgruppe; ITT: Intention-to-Treat; VZ: Verzerrungspotenzial; QoL: gesundheitsbezogene Lebensqualität

a: Die Frauen der Kontrollgruppe und deren Behandler waren für die Diagnose verblindet. Um diese Verblindung möglichst aufrechtzuerhalten, wurden Frauen mit normalen oGTT-Ergebnissen ebenfalls der Kontrollgruppe zugewiesen.

b: siehe Abschnitt „Ergebnisunabhängige Berichterstattung, S. 25

c: Für 8 Geburten der IG und 18 Geburten der KG liegen keine Entbindungsdaten vor. Auch Gründe sind nicht angegeben

d: hoch für perinatale Mortalität, Geburtstrauma, Aufnahme auf Intensivstation und Glukoseinfusionen

e: Inadäquat für den Endpunkt gesundheitsbezogene Lebensqualität (QoL): Von 490 Frauen der IG wurden nur 332 Frauen 6 Wochen nach Einschluss sowie nur 278 Frauen 3 Monate post partum in die QoL-Analyse eingeschlossen. Von 510 in der KG wurden nur 350 Frauen 6 Wochen nach Einschluss und 295 Frauen 3 Monate post partum in die QoL-Analyse eingeschlossen.

f: hoch für den Endpunkt gesundheitsbezogene Lebensqualität

5.2.1.4 Studienpool A: Ergebnisse der randomisierten Studien zu patientenrelevanten Therapiezielen

Die Darstellung der Zielgrößen erfolgt getrennt nach mütterlichen Ergebnissen, nach kindlichen Ergebnissen sowie nach Surrogatparametern. Nur für die in den Methoden definierten patientenrelevanten Endpunkte wird eine zusammenfassende Bewertung zum Nutzen und Schaden abgegeben.

Mütterliche Ergebnisse

5.2.1.4.1 Mütterliche Mortalität

Angaben zur mütterlichen Mortalität werden zwar in keinem der RCTs direkt berichtet, sie lassen sich aber aus Crowther 2005, Langer 1989 und O’Sullivan 1966 ableiten. Da in diesen Studien keine Teilnehmerin vorzeitig aus der Analyse ausgeschlossen wurde, können sich bis Studienende in keiner Gruppe Todesfälle ereignet haben. Für Landon 2009 und Bonomo 2005 kann diese Schlussfolgerung nicht sicher getroffen werden, da ein Teil der randomisierten Schwangeren vorzeitig aus der Analyse ausgeschlossen wurde.

Zusammenfassende Bewertung:

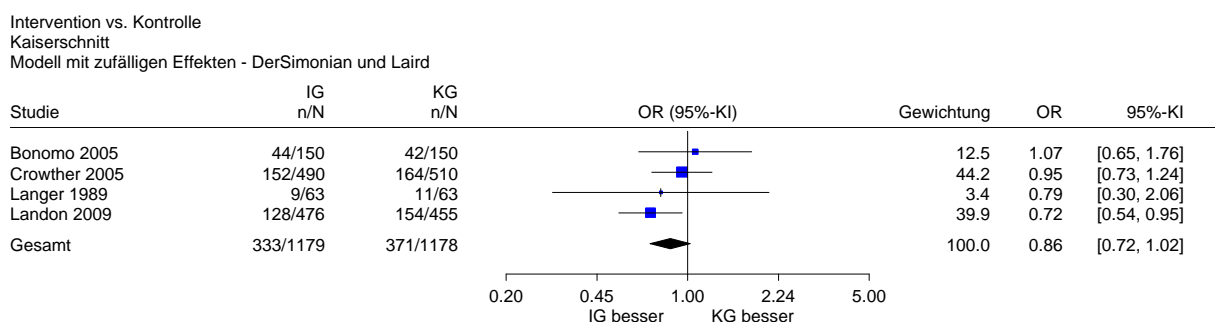
Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

5.2.1.4.2 Art der Geburt

Angaben zum Anteil der Kaiserschnitte waren in 4 Studien vorhanden und sind in Tabelle 10 dargestellt. Nur in Landon 2009 gab es einen statistisch signifikanten Gruppenunterschied.

Die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab hingegen keinen statistisch signifikanten Unterschied (Abbildung 2).



Heterogenität: $Q=2.83$, $df=3$, $p=0.418$, $I^2=0\%$
Gesamteffekt: $Z\text{ Score}=-1.71$, $p=0.087$, $\tau=0$

Abbildung 2: Studienpool A – Meta-Analyse Anteil der Geburten mit Kaiserschnitt

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Tabelle 10: Studienpool A – Kaiserschnittraten

Kaiserschnittraten			
Studie	IG (%)	KG (%)	p-Wert
Landon 2009	128 (27) ^a	154 (34) ^a	0,02
Bonomo 2005	44 (29)	42 (28)	k. A. ^b
Crowther 2005	152 (31) ^c	164 (32)	0,73
Langer 1989	9 (15) ^d	11 (17) ^d	k. A.
O'Sullivan 1966	k. A.	k. A.	k. A.

IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angaben
a: IG: 476; KG: 455
b: Autoren bezeichnen die Ereignisraten in beiden Behandlungsgruppen als vergleichbar
c: Davon hatten 15 % vs. 12 % einen elektiven Kaiserschnitt und 16 % vs. 20 % einen Notfallkaiserschnitt.
d: Summe aus primären und Folgekaiserschnitten

5.2.1.4.3 Andere Geburtsinterventionen

Zu operativen Entbindungen gab es in keiner der 4 Studien Angaben. Angaben zur Einleitung der Geburt werden in Abschnitt 5.2.1.4.21 beschrieben.

5.2.1.4.4 Geburtsverletzungen

Nur Crowther 2005 enthält Angaben zu perinealen Geburtsverletzungen. Danach erlitten 52 % Frauen in der Interventionsgruppe (255 von 490) und 50 % Frauen in der Kontrollgruppe (254 von 510) eine Verletzung. Dieser numerische Unterschied ist nicht statistisch signifikant (adjustiertes RR: 1,05; 95 %-KI [0,93; 1,19]; p = 0,42).

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

5.2.1.4.5 Geburtskomplikationen (zum Beispiel Schulterdystokie, postpartale Blutungen)

Angaben zur Schulterdystokie fanden sich in Landon 2009 und Crowther 2005, zu postpartalen Blutungen und zum Kindbettfieber nur in Crowther 2005.

Schulterdystokien: Tabelle 11 enthält die von Landon 2009 und Crowther 2005 berichteten Ergebnisse zu Schulterdystokien. In Crowther 2005 wurde das Vorliegen und der Schweregrad einer Schulterdystokie vom bei der Geburt anwesenden Geburtshelfer anhand einer standardisierten Checkliste abgeschätzt und dokumentiert [13]. In Landon 2009 überprüften 2 verblindete Reviewer die Diagnose einer Schulterdystokie [6]. In beiden Studien wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingeschätzt.

Eine meta-analytische Zusammenfassung in einem Modell mit zufälligen Effekten ergab einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppen. Das Odds Ratio beträgt 0,40 (95 %-KI [0,21; 0,75]).

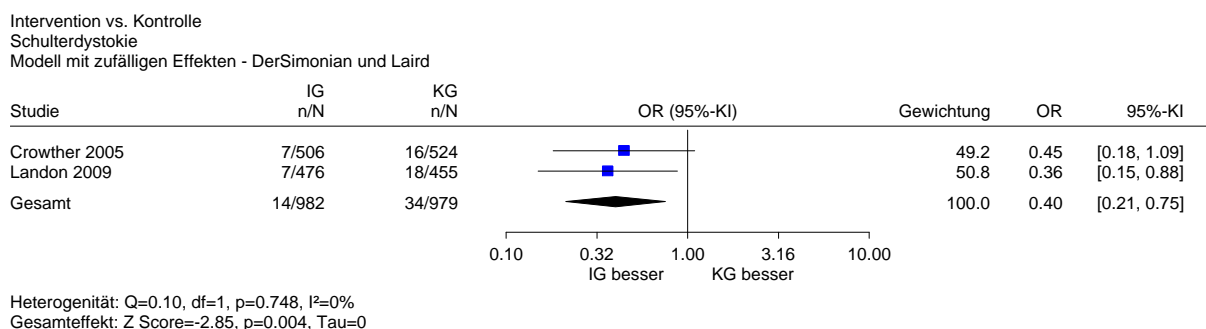


Abbildung 3: Studienpool A – Meta-Analyse der Schulterdystokien

Tabelle 11: Studienpool A – Schulterdystokie

Schulterdystokie				
	IG (%)	KG (%)	RR (95 %-KI)	p-Wert
Landon 2009	7 / 476 (1,5)	18 / 455 (4,0)	0,37 (0,14; 0,97)	0,02
Crowther 2005	7 / 490 (1,4)	16 / 510 (3,1)	0,46 (0,19; 1,10)	0,08

Zusammenfassende Bewertung:
Es liegt ein Beleg für einen Nutzen vor.

Schulterdystokie war in Crowther 2005 zudem noch eine Komponente des prädefinierten kombinierten Endpunkts. Die Bewertung dieses kombinierten Endpunkts ist im Abschlussbericht S07-01 ausführlich diskutiert.

Postpartale Blutungen: Die Häufigkeit postpartaler Blutungen war in Crowther 2005 mit jeweils 6 % in den beiden Gruppen identisch. Kindbettfieber trat bei 3 % der Frauen in der Interventionsgruppe (17 von 490) und 6 % (29 von 510) in der Kontrollgruppe zwar numerisch häufiger auf, der Unterschied war jedoch nicht statistisch signifikant (adjustiertes RR: 0,63; 95 %-KI [0,35; 1,13]; $p = 0,12$).

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

5.2.1.4.6 Präeklampsie / Eklampsie

Der Zielparameter Präeklampsie wurde nur in Landon 2009 berichtet. Die Definition für eine Präeklampsie entspricht in Crowther 2005 einer schwangerschaftsinduzierten Hypertonie, weswegen die publizierten Ereignisse im vorliegenden Bericht diesem Endpunkt zugeordnet wurden.

Präeklampsien traten in Landon 2009 bei 2,5 % der Frauen in der Interventionsgruppe (12 von 476) und 5,5 % (25 von 455) der Frauen in der Kontrollgruppe auf. Der Unterschied war statistisch signifikant (adjustiertes RR: 0,46; 95 %-KI [0,22; 0,97]; $p = 0,02$).

Da Landon 2009 die einzige Studie ist, die zu diesem Endpunkt berichtet, waren keine Meta-Analysen möglich. Das Verzerrungspotenzial des Endpunkts Präeklampsie wurde in Landon 2009 als niedrig eingestuft.

Zusammenfassende Bewertung:

Ein Hinweis auf einen Nutzen liegt vor.

5.2.1.4.7 Dauer des stationären Aufenthalts

Angaben zur Dauer des stationären Aufenthalts fanden sich nur bei Crowther 2005. Für beide Gruppen wurde eine Aufenthaltsdauer von im Median 4 Tagen angegeben (Interquartilsabstand: 3 bis 5 Tage).

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

5.2.1.4.8 Inanspruchnahme medizinischer Leistungen (Untersuchungen, Tests, Behandlungen)

Zu dieser Zielgröße wurden keine Ergebnisse berichtet. Die in Crowther 2005 als Teil der Intervention beobachteten Therapieunterschiede sind in Abschnitt 5.2.1.4.20 beschrieben.

5.2.1.4.9 Psychologische Ergebnisse (zum Beispiel Angst, postnatale Depression)

In Crowther 2005 wurde der Endpunkt Angst mit einem standardisierten Messinstrument zu 2 unterschiedlichen Messzeitpunkten (6 Wochen nach Studieneinschluss und 3 Monate nach der Geburt) erhoben. Es wurde die Kurzform des State Trait Anxiety Inventory (STAI) zur Selbsteinschätzung eingesetzt [22]. Von insgesamt 1000 randomisierten Frauen (IG n = 490; KG n = 510) füllten zu beiden Zeitpunkten weniger als 70 % der Frauen den Bogen aus. Vor der Geburt waren es in der Interventionsgruppe nur 332 Frauen (68 %) und in der Kontrollgruppe 350 Frauen (69 %). Drei Monate nach der Geburt waren es 278 Frauen (57 %) in der Interventionsgruppe und 295 Frauen (58 %) in der Kontrollgruppe. Diese Ausfallzahlen sind so groß, dass die Ergebnisse nicht interpretiert werden können.

Dasselbe gilt für die berichteten Ergebnisse zum Risiko für das Auftreten einer schwangerschaftsassozierten Depression, das mithilfe der Edinburgh Postnatale Depression Scale [23] 3 Monate nach der Geburt bestimmt wurde. An der Befragung nahmen nur 278 Frauen (57 %) in der Interventionsgruppe und 295 Frauen (58 %) in der Kontrollgruppe teil. Auch diese Ergebnisse können wegen der hohen Ausfallrate nicht interpretiert werden.

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

5.2.1.4.10 Gesundheitsbezogene Lebensqualität

In Crowther 2005 wurde die gesundheitsbezogene Lebensqualität mit dem standardisierten Messinstrument SF-36 erhoben [24]. Die Befragung fand zu 2 unterschiedlichen Zeitpunkten statt: vor der Geburt (6 Wochen nach Studieneinschluss) und 3 Monate nach der Geburt. Vor der Geburt hatten von insgesamt 1000 Frauen 682 (68 %) die Fragebögen ausgefüllt, 3 Monate nach der Geburt waren es 573 (57 %).

Die Ergebnisse sind jedoch für die Fragestellung des Berichts nicht interpretierbar. Um die Auswirkungen der Diagnose GDM messen zu können, hätte auch unmittelbar nach Studieneinschluss (das heißt, vor der Mitteilung der Diagnose) eine Befragung stattfinden müssen, um einen eventuellen Effekt der Diagnose zu messen. Sechs Wochen nach Studieneinschluss können diese Auswirkungen bereits abgeklungen sein. Zudem lag die Rate fehlender Angaben nach der Geburt bei 43 %, sodass das Verzerrungspotenzial sehr hoch ist.

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Kindliche Ergebnisse

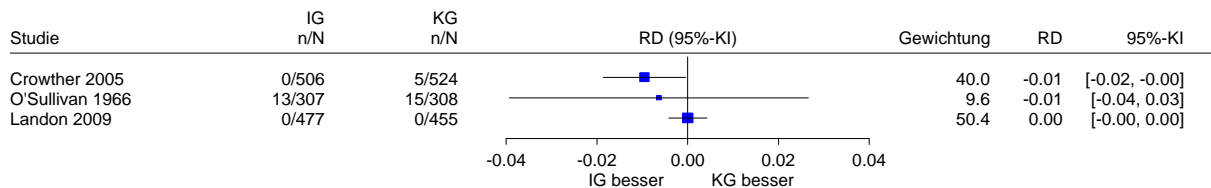
5.2.1.4.11 Perinatale und neonatale Mortalität

Drei der 5 RCTs berichten Angaben zur perinatalen und neonatalen Mortalität, die in Tabelle 12 aufgelistet sind. Auffällig ist in O’Sullivan 1966 die in allen Gruppen sehr hohe Mortalitätsrate. Diese Studie fand unter den vor 40 Jahren üblichen Betreuungsstandards statt, Einflüsse auf die Mortalitätsrate könnten deshalb kaum auf heutige Raten übertragen werden. In Crowther 2005 lag die Rate selbst in der Kontrollgruppe deutlich niedriger als in O’Sullivan 1966. Crowther 2005 und O’Sullivan 1966 berichteten zwar über eine numerisch niedrigere Rate kindlicher Todesfälle in der Interventionsgruppe, nach Berechnungen der Autoren war der Unterschied aber jeweils nicht statistisch signifikant. In Landon 2009 sind keine perinatalen Todesfälle berichtet, allerdings fehlen Daten von 8 Geburten der Interventions- und 18 Geburten der Kontrollgruppe.

Bei den 5 perinatalen Todesfällen in Crowther 2005 handelte es sich um 3 Fehlgeburten und 2 neonatale Todesfälle. Zwei Fehlgeburten waren unerklärte intrauterine Todesfälle bei reifen Kindern normalen Wachstums. Die dritte Fehlgeburt fand in der 35. SSW statt und war assoziiert mit Präeklampsie und einer intrauterinen Wachstumsrestriktion. Ein Kind starb nach der Geburt an einer angeborenen Fehlbildung, das andere Kind starb nach einer Asphyxie während der Geburt.

Bereits im Abschlussbericht S07-01 hatte die Meta-Analyse der perinatalen Todesfälle keinen statistisch signifikanten Unterschied ergeben. Da Landon 2009 keine Ereignisse berichtete, konnte sich keine Änderung der Bewertung ergeben. Da eine Studie ohne Ereignisse in beiden Gruppen zu keiner Veränderung des gepoolten Effektschätzers auf der Basis von Odds Ratios führt, wurde zusätzlich eine meta-analytische Zusammenfassung auf der Basis von Risikodifferenzen durchgeführt. Aus diesem Modell ergab sich jedoch eine hohe Heterogenität (Heterogenitätstest $p < 0,2$), sodass hier kein gepoolter Effektschätzer dargestellt wurde (Abbildung 4).

Intervention vs. Kontrolle
perinatale und neonatale Mortalität
Modell mit zufälligen Effekten - DerSimonian und Laird



Heterogenität: $Q=6.99$, $df=2$, $p=0.030$, $I^2=71.4\%$

Abbildung 4: Studienpool A – Forest-Plot perinatale und neonatale Mortalität

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Die perinatale und neonatale Mortalität war in Crowther 2005 zudem noch eine Komponente des prädefinierten kombinierten Endpunkts. Die Bewertung dieses kombinierten Endpunkts findet sich im Abschlussbericht S07-01.

Tabelle 12: Studienpool A – perinatale und neonatale Mortalität

Studie	perinatale und neonatale Mortalität (IG vs. KG)			
	gesamt	neonataler Tod (bis 28. Tag nach Geburt)	Totgeburt	perinataler Tod (> 28. SSW bis 1. Woche nach Geburt)
Landon 2009	0 vs. 0	0 vs. 0	0 vs. 0	0 vs. 0
Bonomo 2005	k. A.	k. A.	k. A.	k. A.
Crowther 2005	0 vs. 5 (1 %) p = 0,07	0 vs. 2 (< 1 %)	0 vs. 3 (1 %)	k. A.
Langer 1989	k. A.	k. A.	k. A.	k. A.
O’Sullivan 1966	13 (4,3 %) ^a vs. 15 (4,9 %) ^a ; ns ^b	13 (4,3 %) ^a vs. 15 (4,9 %) ^a ; ns ^b	8 (2,6 %) vs. 8 (2,6 %) ^c ; ns ^b	13 (4,3 %) vs. 15 (4,9 %) ^a ; ns ^b
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe; ns: nicht statistisch signifikant a: Summe der Todesfälle von lebensfähigen Feten nach der 28. SSW und der neonatalen Todesfälle bis zum 14. Tag nach der Geburt b: In der Publikation wurde kein p-Wert angegeben. c: Tod von lebensfähigen Feten nach der 28. SSW				

5.2.1.4.12 Anteil der Fehlgeburten

Zu dieser Zielgröße gab es in keiner der 5 Studien Angaben.

5.2.1.4.13 Geburtstraumata (zum Beispiel Verletzungen des Brachialplexus, Knochenbrüche)

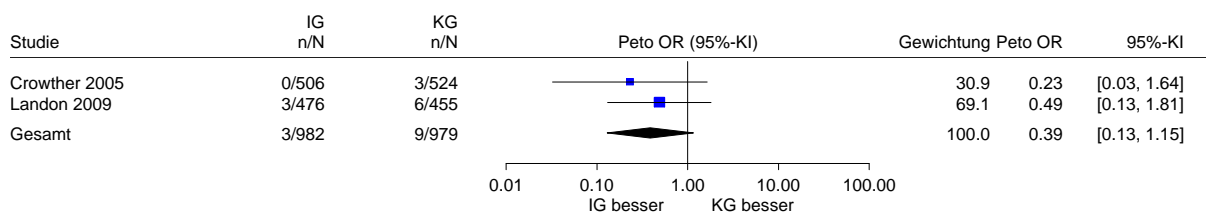
Lediglich Landon 2009 und Crowther 2005 berichten über Geburtstraumata. Tabelle 13 enthält die Ergebnisse der beiden Studien.

Eine meta-analytische Zusammenfassung in einem Modell mit Peto Odds Ratios ergab keinen statistisch signifikanten Unterschied (Abbildung 5).

Intervention vs. Kontrolle

Geburtstraumata

Modell mit festem Effekt - Peto Odds Ratio



Heterogenität: $Q=0.39$, $df=1$, $p=0.533$, $I^2=0\%$

Gesamteffekt: Z Score=-1.71, $p=0.088$

Abbildung 5: Studienpool A – Meta-Analyse der Geburtstraumata

Tabelle 13: Studienpool A – Geburtstraumata

Geburtstraumata				
	IG (%)	KG (%)	RR (95 %-KI)	p-Wert
Landon 2009	3 / 476 (0,6)	6 / 455 (1,3)	0,48 (0,10; 2,20)	0,33
Crowther 2005	0 / 506 (0)	3 / 524 (0,6)		0,11

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Geburtstraumata waren in Crowther 2005 zudem noch eine Komponente des prädefinierten kombinierten Endpunkts. Die Bewertung dieses kombinierten Endpunkts findet sich im Abschlussbericht S07-01.

5.2.1.4.14 Diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen (zum Beispiel Glukoseinfusionen aufgrund von Hypoglykämie, Fototherapie aufgrund von Hyperbilirubinämie, Beatmung aufgrund eines akuten Atemnotsyndroms) inkl. Aufnahme auf eine Intensivstation und Dauer des stationären Aufenthalts

Die Angaben zu der Zielgröße diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen, sind in Tabelle 14 dargestellt.

Fototherapie aufgrund von Hyperbilirubinämie: Bezüglich des Endpunkts Hyperbilirubinämie berichtete nur Crowther 2005 über den Bedarf an einer Fototherapie. Der Anteil der Kinder unterschied sich nicht zwischen den Gruppen.

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Glukoseinfusionen aufgrund von Hypoglykämie: Zum Endpunkt Hypoglykämie berichteten lediglich Landon 2009 und Crowther 2005 über Ereignisse, bei denen intravenös Glukose gegeben wurde.

Eine meta-analytische Zusammenfassung ergab jedoch eine hohe Heterogenität mit einander entgegengesetzten Effektschätzern (Heterogenitätstest $p < 0,2$), sodass hier kein gepoolter Effektschätzer dargestellt wurde (Abbildung 6).

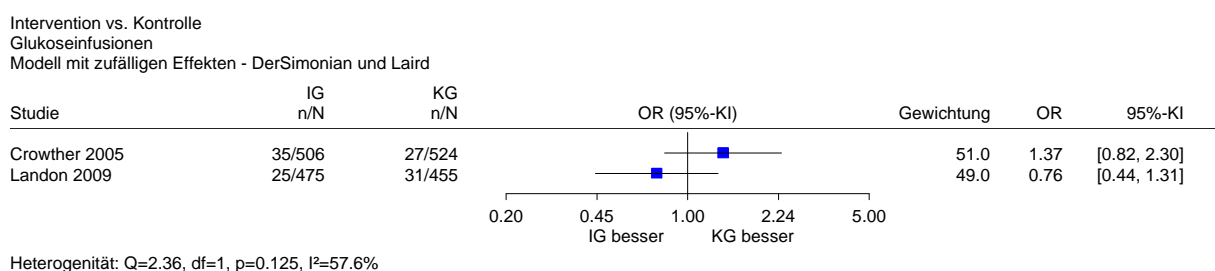


Abbildung 6: Studienpool A – Meta-Analyse der Glukoseinfusionen

Zusammenfassende Bewertung:

Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Beatmung aufgrund eines akuten Atemnotsyndroms: Die Zahl der Beatmungen aufgrund eines akuten Atemnotsyndroms wurde nur in Crowther 2005 berichtet (siehe Tabelle 14). Der Unterschied war nicht statistisch signifikant.

Zusammenfassende Bewertung:

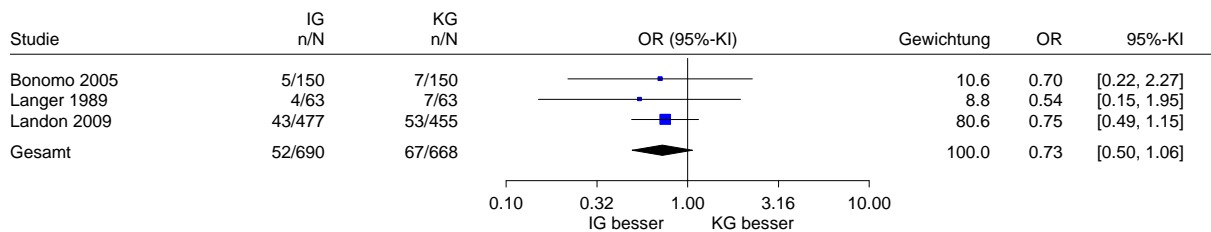
Kein Beleg für und kein Hinweis auf einen Nutzen liegt vor.

Kein Beleg für und kein Hinweis auf einen Schaden liegt vor.

Aufnahme auf eine Intensivstation: Der Endpunkt Aufnahme auf eine Intensivstation wurde in 3 Studien berichtet. In Bonomo 2005 wurden 3 bis 5 %, in Langer 1989 6 bis 11 % und in Landon 2009 9 bis 12 % der Kinder zur neonatalen Intensivversorgung („neonatal intensive care“) überwiesen. Der Unterschied war in keiner der 3 Studien zwischen den Gruppen statistisch signifikant.

Auch die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab keinen statistisch signifikanten Unterschied (Abbildung 7).

Intervention vs. Kontrolle
neonatale Intensivpflege
Modell mit zufälligen Effekten - DerSimonian und Laird



Heterogenität: $Q=0.23$, $df=2$, $p=0.893$, $I^2=0\%$
Gesamteffekt: Z Score=-1.65, $p=0.098$, $\tau=0$

Abbildung 7: Studienpool A – Meta-Analyse Aufnahme auf eine Intensivstation

Crowther 2005 berichtete über die Aufnahme auf eine „neonatale Pflegestation“ („neonatal nursery“). Die Indikation zur Überweisung eines Kindes wurde in dieser Studie völlig anders als in Bonomo 2005 und in Langer 1989 gestellt, offenbar auch schon bei leichten Auffälligkeiten. Die Indikationen zur Überweisung wurden nicht berichtet. In Crowther 2005 wurden 71 % der Kinder der Interventionsgruppe und 61 % der Kinder in der Kontrollgruppe auf eine solche neonatale Pflegestation verlegt. Die Ergebnisse können deshalb nicht als Aufnahme auf eine Intensivstation im Sinne des Berichtsplans bewertet werden und wurden folglich in der Meta-Analyse nicht berücksichtigt. Der Anteil war in der Interventionsgruppe im Vergleich zur Kontrollgruppe statistisch signifikant erhöht. Die in der Interventionsgruppe erhöhte Zahl der Aufnahmen auf eine neonatale Pflegestation in Crowther 2005 wird als Element der Intervention bewertet.

Zusammenfassende Bewertung:

Es liegt kein Beleg für oder Hinweis auf einen Nutzen vor.

Es liegt kein Beleg für oder Hinweis auf einen Schaden vor.

Dauer des stationären Aufenthalts: Zur Zielgröße Dauer des stationären Aufenthalts gab es in keinem der RCTs Angaben.

Tabelle 14: Studienpool A – diagnostische und therapeutische Maßnahmen, die über das übliche Maß hinausgehen

Studie	Foto- therapie	Glukose- infusionen	Beatmung aufgrund Atemnot- syndrom	Aufnahme Intensivstation	Dauer stationärer Aufenthalt
Landon 2009	k. A.	25 / 475 (5,3) 31 / 455 (6,8) p = 0,32	k. A.	43 / 477 (9,0 %) 53 / 455 (11,6 %) p = 0,19	k. A.
Bonomo 2005	k. A.	k. A.	k. A.	3,3 % vs. 4,7 %; ns	k. A.
Crowther 2005	IG: 44 (9 %) vs. KG: 48 (9 %) ^a ; p = 0,72	35 (7 %) vs. 27 (5 %) ^b ; p = 0,16	27 (5 %) vs. 19 (4 %) ^c ; p = 0,15	k. A. ^d	k. A.
Langer 1989	k. A.	k. A.	k. A. ^e	4 (6 %) vs. 7 (11 %); ns	k. A.
O'Sullivan 1966	k. A.	k. A.	k. A.	k. A.	k. A.
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe; ns: nicht signifikant, keine p-Werte publiziert a: Hyperbilirubinämie mit der Notwendigkeit einer Fototherapie b: Hypoglykämie mit der Notwendigkeit der Gabe von Glukose intravenös, wobei der Hypoglykämiewert, bei dem eine Therapie erforderlich war, vom Kliniker festgelegt wurde c: Bedarf an zusätzlicher Sauerstoffgabe innerhalb der ersten 4 Stunden nach Geburt, verabreicht auf der neonatalen Pflegestation d: In Crowther 2005 wurde nur „Aufnahme auf neonatale Pflegestation“ berichtet: IG: 357 (71 %) vs. KG: 321 (61 %); p = 0,01 e: Angaben beziehen sich auf die gesamte Studienpopulation mit 3,2 %					

5.2.1.4.15 Unerwünschte Folgen einer Intervention mit blutzuckersenkender Intention oder anderer Interventionen, die auf eine Optimierung des geburtshilflichen Managements ausgerichtet sind

Zu diesen Zielgrößen gab es in keiner der 5 RCTs Angaben.

Surrogatparameter

5.2.1.4.16 Geburtsgewicht / Makrosomie / LGA / SGA

In Landon 2009, Bonomo 2005, Crowther 2005 und Langer 1989 fanden sich Zahlenangaben zum mittleren Geburtsgewicht, wobei sich nur in Landon 2009 und Crowther 2005 ein statistisch signifikanter Unterschied zugunsten der Interventionsgruppe zeigte, während der Unterschied in den beiden anderen Studien nicht statistisch signifikant war. Für O’Sullivan 1966 kann das mittlere Geburtsgewicht nur ungefähr anhand einer Abbildung geschätzt werden – Näheres siehe Tabelle 15.

Makrosomie: Die Ergebnisse bleiben prinzipiell auch bei Betrachtung der aus dem Geburtsgewicht resultierenden Kategorie „Makrosomie“ unverändert. Hier war in Landon 2009, Crowther 2005 und O’Sullivan 1966 ein statistisch signifikanter Unterschied zugunsten der Interventionsgruppe vorhanden (Tabelle 15).

Die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppen (Abbildung 8). Das Odds Ratio betrug 0,38 (95 %-KI 0,30; 0,49).

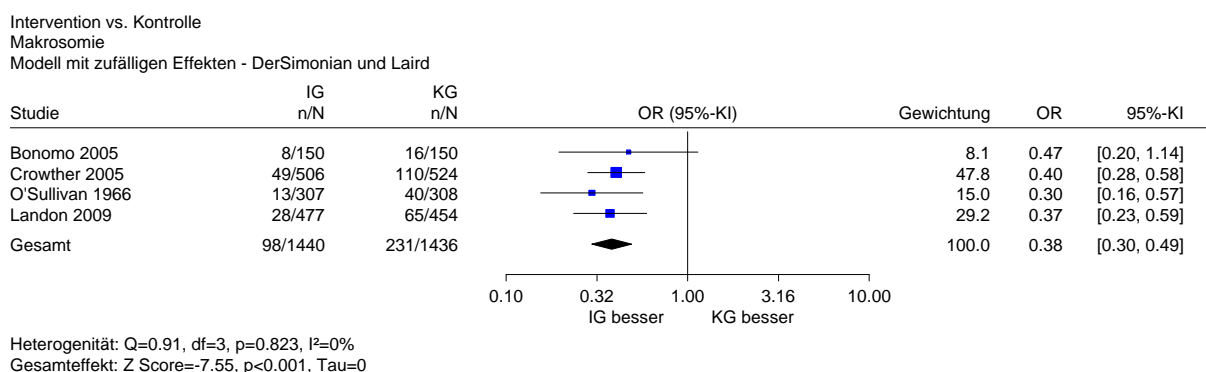


Abbildung 8: Studienpool A – Meta-Analyse Makrosomie

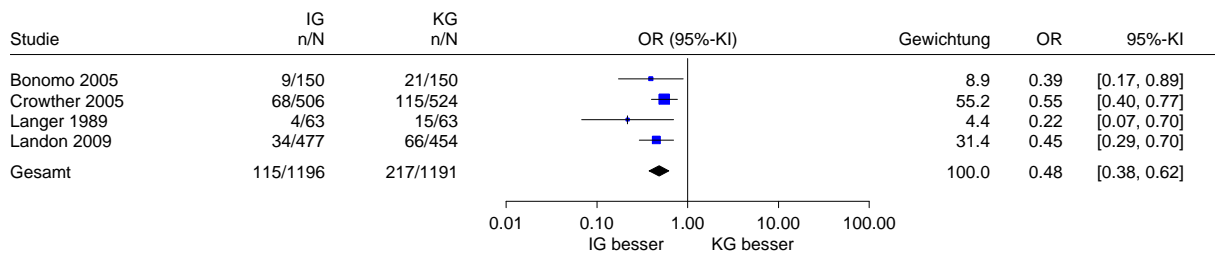
LGA: In Landon 2009, Bonomo 2005, Crowther 2005 und Langer 1989 fand sich für den Endpunkt Large-for-Gestational-Age (LGA) ein statistisch signifikanter Unterschied zugunsten der Interventionsgruppe. In O’Sullivan 1966 wird über den Endpunkt nicht berichtet.

Die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppen (Abbildung 9). Das Odds Ratio betrug 0,48 (95 %-KI 0,38; 0,62).

Intervention vs. Kontrolle

LGA

Modell mit zufälligen Effekten - DerSimonian und Laird



Heterogenität: $Q=2.79$, $df=3$, $p=0.425$, $I^2=0\%$
Gesamteffekt: Z Score=-5.85, $p<0.001$, Tau=0

Abbildung 9: Studienpool A – Meta-Analyse Large-for-Gestational-Age

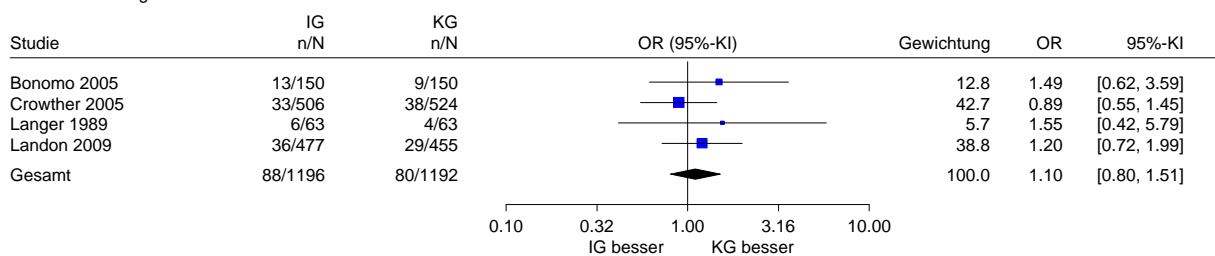
SGA: Landon 2009, Bonomo 2005, Crowther 2005 und Langer 1989 machen Angaben zur Häufigkeit von Kindern Small-for-Gestational-Age (SGA). Es gab keinen statistisch signifikanten Unterschied. O’Sullivan 1966 enthält keine Angaben zu SGA (Tabelle 15).

Die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab keinen statistisch signifikanten Unterschied (Abbildung 10).

Intervention vs. Kontrolle

SGA

Modell mit zufälligen Effekten - DerSimonian und Laird



Heterogenität: $Q=1.54$, $df=3$, $p=0.672$, $I^2=0\%$
Gesamteffekt: Z Score=0.61, $p=0.543$, Tau=0

Abbildung 10: Studienpool A – Meta-Analyse Small-for-Gestational-Age

Tabelle 15: Studienpool A – Geburtsgewicht und verwandte Endpunkte

Studie	Angaben zum Geburtsgewicht			
	Geburtsgewicht ^a (g)	LGA (≥ 90. Perzentile)	Makrosomie (≥ 4000 g)	SGA (≤ 10. Perzentile)
Landon 2009	IG: 3302 (502) KG 3408 (589) p < 0,001	7,1 % (34 / 477) 14,5 % (66 / 454) p < 0,001	5,9 % (28 / 477) 14,3 % (65 / 454) p < 0,001	7,5 % (36 / 477) 6,4 % (29 / 455) p = 0,49
Bonomo 2005	IG: 3365 (436) vs. KG: 3437 (462); ns	6,0 % vs. 14,0 %; p = 0,046	5,3 % vs. 10,7 %; ns	8,7 % vs. 6,0 %; ns
Crowther 2005	IG: 3335 (551) vs. KG: 3482 (660); p < 0,001	13 % (68) vs. 22 % (115); p < 0,001	10 % (49) vs. 21 % (110); p < 0,001	7 % (33) vs. 7 % (38); p = 0,59
Langer 1989	IG: 3261 (496) vs. KG: 3422 (584); ns	6 % (4) vs. 24 % (15); p < 0,03	k. A.	10 % (6) vs. 6 % (4); ns
O'Sullivan 1966	IG: k. A. ^b KG:	k. A.	4,3 % (13) vs. 13,1 % (40) ^c ; p = k. A.	k. A.
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe; ns: nicht signifikant, keine p-Werte publiziert; LGA: Large-for-Gestational-Age; SGA: Small-for-Gestational-Age a: Mittelwerte, ggf. gerundet, mit Angabe der Standardabweichung in Klammern b: Nur grafische Darstellung publiziert. Aus Abbildung 1 kann das Geburtsgewicht mit etwa 3200 g (IG) vs. 3500 g (KG) geschätzt werden. c: Makrosomie war definiert als ≥ 4100 g in der 40. SSW.				

5.2.1.4.17 Apgar-Wert

Angaben zum Apgar-Wert fanden sich in Bonomo 2005 und Crowther 2005 (Tabelle 16). In keiner der Studien zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Tabelle 16: Studienpool A – Apgar-Werte

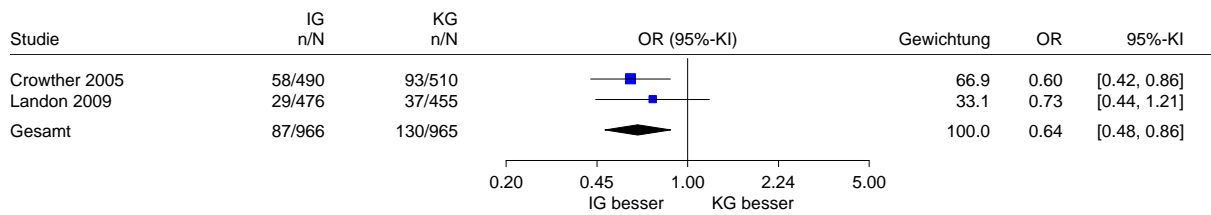
Studie	Apgar-Wert
Landon 2009	k. A. IG: KG:
Bonomo 2005	IG: 9,7 (0,5) vs. KG: 9,5 (0,5) ^a ; ns
Crowther 2005	IG: 6 (1 %) vs. KG: 11 (2 %) ^b ; p = 0,26
Langer 1989	IG: k. A. KG:
O’Sullivan 1966	IG: k. A. KG:
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe; ns: nicht signifikant, keine p-Werte publiziert a: Apgar-Wert nach 5 Minuten b: Kinder mit einem 5-Minuten-Apgar-Wert < 7	

5.2.1.4.18 Schwangerschaftsbedingte Hypertonie

Der Endpunkt schwangerschaftsbedingte Hypertonie wurde in Landon 2009 und in Crowther 2005 berichtet. In Crowther wird er als „Präeklampsie“ bezeichnet, definiert als zweimalig gemessener Blutdruck $\geq 140/90$ mmHg. In Landon 2009 lassen sich die Angaben zu schwangerschaftsbedingter Hypertonie (zweimalig gemessener Blutdruck $\geq 140/90$ mmHg oder einmalig erhöhter Blutdruck mit Einleitung einer medikamentösen Therapie) aus Tabelle 4 der Publikation [6] errechnen.

Die meta-analytische Zusammenfassung der Studien in einem Modell mit zufälligen Effekten ergab einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppen (Abbildung 11). Das Odds Ratio betrug 0,64 (95 %-KI 0,48; 0,86).

Intervention vs. Kontrolle
Schwangerschaftsbedingte Hypertonie
Modell mit zufälligen Effekten - DerSimonian und Laird



Heterogenität: $Q=0.39$, $df=1$, $p=0.531$, $I^2=0\%$
Gesamteffekt: Z Score=-2.99, $p=0.003$, Tau=0

Abbildung 11: Studienpool A – Meta-Analyse schwangerschaftsbedingte Hypertonie

Tabelle 17: Studienpool A – Schwangerschaftsbedingte Hypertonie

Studie	Schwangerschaftsbedingte Hypertonie
Landon 2009 ^a	IG: 6,1 % (29 / 476) KG: 8,1 % (37 / 455) p = 0,23 ^b
Bonomo 2005	k. A. IG: KG:
Crowther 2005	IG: 12 % (58 von 490) KG: 18 % (93 von 510) p = 0,02
Langer 1989	k. A. IG: KG:
O'Sullivan 1966	k. A. IG: KG:
IG: Interventionsgruppe; KG: Kontrollgruppe. k. A.: keine Angabe; ns: nicht signifikant, keine p-Werte publiziert a: Differenz aus dem kombinierten Endpunkt Präeklampsie oder Schwangerschaftshypertonie und dem Endpunkt Präeklampsie Tabelle 4 [6] b: eigene Berechnung; unbedingter exakter Test (CSZ-Methode [25])	

5.2.1.4.19 Spätere Diagnose eines Diabetes mellitus der Mutter

Angaben zu diesem Endpunkt fanden sich in der Studie von O'Sullivan 1966. Innerhalb von 16 Jahren nach der Diagnose GDM entwickelten in der Interventionsgruppe 34,9 % der Frauen einen manifesten Diabetes mellitus, in der Kontrollgruppe waren es 35,7 % der Frauen. Der Unterschied war nicht statistisch signifikant.

Ergänzende Ergebnisse

5.2.1.4.20 Inanspruchnahme medizinischer Leistungen (Untersuchungen, Tests, Behandlungen)

Ausführliche Angaben zur Inanspruchnahme medizinischer Leistungen fanden sich in Crowther 2005. In Landon 2009 findet sich nur die Angabe, dass Frauen der Interventionsgruppe im Durchschnitt 7 pränatale Visiten in Anspruch nahmen, bei Frauen der Kontrollgruppe waren es 5 ($p < 0,001$). In der Interventionsgruppe wurden 37 Frauen mit Insulin behandelt, in der Kontrollgruppe waren es 2.

Die berichteten Unterschiede sind beabsichtigter Teil der Intervention und daher nicht als Nutzen oder Schaden zu bewerten. Zur Bewertung der Aspekte „Einleitung der Geburt“ und „Aufnahme auf eine neonatale Pflegestation“ siehe jeweils die Diskussion in den eigenen Abschnitten.

In Crowther 2005 wurden für diese Analyse des Aufwands in einer eigenen Publikation nur die Einlingsschwangerschaften ausgewertet (siehe Tabelle 18). Die Zahl der vorgeburtlichen Klinikeinweisungen unterschied sich nicht statistisch signifikant zwischen den Gruppen. Interventionsbedingt verlagerten sich die Kontakte von der „antenatal clinic“ hin zu Spezialisten, Diätberatern und zu Besuchen einer Diabetesschulung; auch andere ärztliche Konsultationen waren häufiger. In der Interventionsgruppe wurden 20,3 % der Frauen mit Insulin behandelt im Vergleich zu 3,2 % der Frauen in der Kontrollgruppe. Die Dauer der Fototherapie aufgrund von Hyperbilirubinämie war nicht statistisch signifikant unterschiedlich; auch die Anzahl der Geburten, bei denen ein Kinderarzt anwesend war, unterschied sich nicht.

Tabelle 18: Crowther 2005 – Inanspruchnahme medizinischer Leistungen nach Aufnahme in die Studie (nur Einlingsschwangerschaften)

Medizinische Leistung	IG n = 474 Mittelwert (IQR)	KG n = 496 Mittelwert (IQR)	adjustierter^a Effekt mittlere Differenz (95 %-KI)	p-Wert
vorgeburtliche Betreuung				
Klinikeinweisungen	135 (28,5 %)	133 (26,8 %)	1,11 (0,91; 1,36)	0,31
Besuche Schwangerensprechstunde	4,4 (1;7)	5,3 (3;7)	-0,70 (-1,06; -0,33)	< 0,001
ärztliche Konsultationen physician clinics ^b	3 (1;7) ^c	0 (0,2) ^c		< 0,001
Besuche Spezialist	4,0 (1;7)	1,3 (0;2)	2,50 (2,13;2,87)	< 0,001
Besuche Diätberater	1,7 (1;2)	0,2 (0;0)	1,56 (1,39;1,72)	< 0,001
Besuche Diabetesschulung	2,0 (1;2)	0,2 (0;0)	1,79 (1,59;1,98)	< 0,001
Insulintherapie	96 (20,3 %)	16 (3,2 %)	6,18 (3,69;10,35)	< 0,001
Betreuung während der Geburt				
Einleitung der Geburt	183 (38,6 %)	148 (29,8 %)	1,34 (1,13;1,60)	0,001
Kaiserschnitt	142 (30,0 %)	153 (30,8 %)	0,97 (0,80;1,17)	0,76
neonatale Betreuung				
Kinderarzt bei Geburt anwesend	211 (44,5 %)	225 (45,4 %)	1,02 (0,89;1,17)	0,72
Aufnahme auf neonatale Pflegestation	330 (69,6 %)	294 (59,3 %)	1,15 (1,04;1,26)	0,004
mittlere Dauer der Fototherapie pro Kind (h)	45 (24;72)	36 (24;51)		0,27
IG: Interventionsgruppe; KG: Kontrollgruppe; IQR: Interquartilsabstand a: adjustiert für mütterliches Alter, Ethnie und Parität b: laut Tabelle 5, Crowther 2005 c: Median (Interquartile); inklusive Mehrlingsschwangerschaften IG: n = 490; KG n = 510				

5.2.1.4.21 Einleitung der Geburt

Landon 2009 und Crowther 2005 enthalten Angaben zur Einleitung der Geburt (siehe Tabelle 19). Der in der Interventionsgruppe von Crowther 2005 beobachtete erhöhte Anteil der Frauen mit Einleitung der Geburt wird als Element der Intervention bewertet und nicht als Schaden.

Eine meta-analytische Zusammenfassung ergab eine hohe Heterogenität (Heterogenitätstest $p < 0,2$), sodass hier kein gepoolter Effektschätzer dargestellt wurde (Abbildung 12). Eine Erklärung für diese Heterogenität ist nicht zu erkennen.

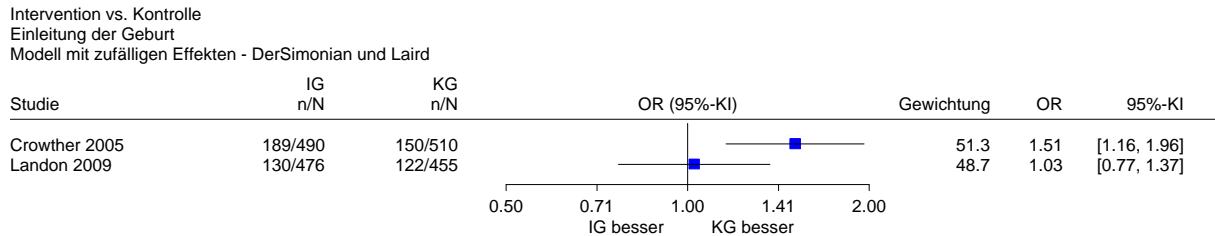


Abbildung 12: Studienpool A – Meta-Analyse Einleitung der Geburt

Tabelle 19: Studienpool A – Einleitung der Geburt

Studie	Einleitung der Geburt
Landon 2009	IG: 27,3 % (130 / 476) KG: 26,8 % (122 / 455) $p = 0,86$
Bonomo 2005	k. A. IG: KG:
Crowther 2005	IG: 39 % (189 / 490) KG: 29 % (150 / 510) $p = 0,02$
Langer 1989	k. A. IG: KG:
O’Sullivan 1966	k. A. IG: KG:
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe	

5.2.1.4.22 Ergebnisse zu Laborparametern

Hyperbilirubinämie: Bezüglich des Endpunkts Hyperbilirubinämie berichtete nur Crowther 2005 über den Bedarf an einer Fototherapie (Tabelle 14). Der Anteil der Kinder unterschied sich nicht zwischen den Gruppen. Landon 2009, Langer 1989 und Bonomo 2005 berichteten

lediglich über den Anteil der Kinder mit erhöhten Bilirubinwerten (Tabelle 20); auch hier gab es aber keinen statistisch signifikanten Unterschied zwischen den Gruppen. Auf eine Meta-Analyse wurde verzichtet, da in den Studien Hyperbilirubinämie unterschiedlich definiert war.

Hypoglykämie: Zum Endpunkt Hypoglykämie berichteten lediglich Landon 2009 und Crowther 2005 über Ereignisse, bei denen intravenös Glukose gegeben wurde (Tabelle 14). Der Anteil lag in der Interventionsgruppe numerisch höher, der Unterschied war jedoch nicht statistisch signifikant. In Bonomo 2005 und Langer 1989 wurden Hypoglykämien nur anhand der Blutglukosewerte der Neugeborenen definiert (Tabelle 20). Langer 1989 berichtete im Unterschied zu Crowther 2005 über eine statistisch signifikante Reduktion der Hypoglykämien in der Interventionsgruppe. Aufgrund der Definition bleibt die Patientenrelevanz der Ereignisse unklar. Auf eine Meta-Analyse wurde verzichtet, da in den Studien Hypoglykämie unterschiedlich definiert war.

Tabelle 20: Studienpool A – Ergebnisse zu Laborparametern

Studie	Hyperbilirubinämie	Hypoglykämie
Landon 2009	IG: 9,6 % (43 / 450) ^a KG: 12,9 % (54 / 418) p = 0,12	16,3 % (62 / 381) ^b 15,4 % (55 / 357) p = 0,75
Bonomo 2005	IG: 4 % vs. KG: 2,7 % ^c ; ns	3,3 % vs. 4 % ^d ; ns
Langer 1989	IG: 4 (6 %) vs. KG: 9 (14 %) ^e ; ns	1 (2 %) vs. 8 (13 %) ^b ; p < 0,02
O’Sullivan 1966	k. A. IG: KG:	k. A.
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angabe; ns: nicht signifikant, keine p-Werte publiziert a: Bilirubin > 95. Perzentil, 16 bis 36 h nach Geburt b: Blutglukose < 35 mg/dl c: Bilirubin ≥ 250 µmol/l (entspricht > 14,6 mg/dl) d: Blutglukose < 31 mg/dl e: Bilirubin > 12 mg/dl		

5.2.1.4.23 Angaben zum Gestationsalter bei Geburt und Frühgeburtlichkeit

In 4 der 5 RCTs fanden sich Angaben zur mittleren oder medianen SSW bei Geburt. Der Unterschied zwischen den Vergleichsgruppen war in Landon 2009, Bonomo 2005 und Langer 1989 als nicht statistisch signifikant beschrieben, in Crowther 2005 fehlten entsprechende

Angaben. In O’Sullivan 1966 wurden keine Daten zur SSW bei Geburt publiziert (Tabelle 21).

Landon 2009, Langer 1989 und O’Sullivan 1966 berichteten über die Häufigkeit von Frühgeburten in den Vergleichsgruppen, wobei der Unterschied nicht statistisch signifikant war (Tabelle 21).

Tabelle 21: Studienpool A – Gestationsalter bei Geburt und Frühgeburtenrate

Studie	Gestationsalter bei Geburt (SSW) ^a	Frühgeburtenrate (< 37. SSW)
Landon 2009	IG: 39,0 (1,8) KG: 38,9 (1,8) p = 0,87	9,4 % (45 / 477) 11,6 % (53 / 455) p = 0,27
Bonomo 2005	IG: 39,4 (1,2) vs. KG: 39,6 (1,7); ns	k. A.
Crowther 2005	IG: 39,0 vs. KG: 39,3 ^b ; p = k. A.	k. A.
Langer 1989	IG: 39 (2) vs. KG: 39 (1); ns	6 % (4) vs. 6 % (4); ns
O’Sullivan 1966	k. A. IG: KG:	8,5 % (26) vs. 7,8 % (24); ns
IG: Interventionsgruppe; KG: Kontrollgruppe; k. A.: keine Angaben; SSW: Schwangerschaftswoche; ns: nicht signifikant, keine p-Werte publiziert a: Mittelwerte (Standardabweichung) b: Median		

5.2.1.4.24 Spätere Entwicklung von Übergewicht des Kindes

Zu diesem Endpunkt gab es in keinem der 5 RCTs Angaben.

5.2.1.5 Landon 2009: Bewertung des kombinierten Endpunkts

In Landon 2009 wurde als primärer Endpunkt eine Kombination von Ereignissen verwendet (siehe Tabelle 22). Für den primären Endpunkt gab es keinen statistisch signifikanten Unterschied zwischen den Gruppen. Dieser kombinierte Endpunkt wird nicht als valide bewertet, da die Aspekte Hypoglykämie, Hyperbilirubinämie und Nabelschnur-C-Peptid keine patientenrelevanten Ergebnisse darstellen.

Tabelle 22: Landon 2009 – Ergebnisse des primären kombinierten Endpunkts

Endpunkt	IG	KG	relatives Risiko (97 %-KI)	p-Wert
N randomisiert	485	473		
Kombinierter Endpunkt	149 / 460	163 / 440	0,87 (0,72; 1,07)	0,14
- Hypoglykämie ^a	62 / 381	55 / 357	1,06 (0,73; 1,53)	0,75
- Hyperbilirubinämie ^b	43 / 450	54 / 418	0,74 (0,49; 1,12)	0,12
- Nabelschnur-C-Peptid ^c	75 / 423	92 / 403	0,78 (0,57; 1,05)	0,07
- Fehlgeburt oder neonataler Tod	0 / 476	0 / 455		
- Geburtstraumata ^d	3 / 476	6 / 455	0,48 (0,10; 2,20)	0,33

IG: Interventionsgruppe; KG: Kontrollgruppe
a: Blutglukose < 35 mg/dl innerhalb von 2 h nach der Geburt
b: > der 95. Perzentile, 16 bis 36 h nach Geburt
c: > der 95. Perzentile (> 1,77 ng/ml)
d: Verletzungen des Brachialplexus oder Schlüsselbein-, Oberarm-, Schädelfrakturen

5.2.1.6 Studienpool A: Zusammenfassung der Ergebnisse

Aus den 5 RCTs des aktualisierten Studienpools A ergibt sich folgendes Bild zum Vergleich zwischen GDM-spezifischen Therapien und einer Routinebetreuung:

Die Studien liefern keinen Beleg dafür, dass durch eine GDM-spezifische Therapie Risiken für das Leben der Mutter oder des Kindes reduziert werden.

Landon 2009 und Crowther 2005 haben unter den Studien des Pools A das größte Gewicht. Mit jeweils etwa 1000 Teilnehmerinnen handelt es sich um die bei Weitem größten RCTs; die Studien berichten Ergebnisse zu vielen Endpunkten und haben aufgrund der methodischen Eigenschaften die im Vergleich höchste Ergebnissicherheit.

Nach Ergänzung der Ergebnisse aus Landon 2009 lassen sich auf Basis von Studienpool A auch zu einzelnen Endpunkten Aussagen machen.

Nutzen:

In der Gesamtabwägung wurde Crowther 2005 weiterhin als Hinweis auf einen Nutzen hinsichtlich „ernster perinataler Komplikationen“ bewertet. Die Größe des Nutzens blieb jedoch unklar.

Für den Endpunkt Schulterdystokie liegt ein Beleg für einen Nutzen durch eine GDM-spezifische Therapie vor.

Für den Endpunkt Präeklampsie liegt ein Hinweis auf einen Nutzen vor. Dieser Hinweis basiert auf Landon 2009.

Schaden:

Für keinen patientenrelevanten Endpunkt liegt ein Beleg für oder ein Hinweis auf einen Schaden durch eine GDM-spezifische Therapie vor.

In Crowther 2005 waren die Raten für die Ereignisse Einleitung der Geburt und Überweisungen auf eine neonatale Pflegestation in der Interventionsgruppe gegenüber der Kontrollgruppe erhöht. Diese erhöhten Raten wurden als Elemente der GDM-spezifischen Intervention bewertet und nicht als Schaden.

Surrogatparameter

Ein einheitliches Ergebnis der Studien in Pool A ist, dass eine GDM-spezifische Therapie zu einem geringeren Anteil makrosomer Kinder führte. Die Häufigkeit dieses Surrogatendpunkts war in allen Studien in den Interventionsgruppen mindestens halbiert, die Reduktion erreichte auch in einzelnen Studien statistische Signifikanz.

Die Meta-Analysen ergaben sowohl für Makrosomie als auch für LGA statistisch signifikante niedrigere Raten in den Interventionsgruppen. Die Odds Ratios betragen 0,38 und 0,48. Das ist ein Nachweis, dass eine GDM-spezifische Therapie den Anteil der Kinder über 4000 g oder über der 90. Perzentile reduzieren kann.

Die Meta-Analyse für SGA ergab ein Odds Ratio von 1,10 zuungunsten der Interventionsgruppen. Diese Erhöhung war nicht statistisch signifikant. Dieses Ergebnis spricht gegen die Befürchtung, dass an die Stärke der Glukosetoleranzstörung angepasste Therapien, die auf eine Reduktion der Makrosomie- / LGA-Raten zielen, auch deutlich die SGA-Raten erhöhen.

Die Meta-Analyse für schwangerschaftsbedingte Hypertonie ergab einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppen mit einem Odds Ratio von 0,64.

5.3 Zusammenfassung der Änderungen aufgrund der Aktualisierung

Im Abschlussbericht S07-01 basierten die Schlussfolgerungen vor allem auf der Bewertung des kombinierten Endpunkts aus Crowther 2005. Aus dem kombinierten Endpunkt war ein Hinweis auf einen Nutzen einer GDM-spezifischen Therapie abgeleitet worden.

Der Einschluss von Landon 2009 in Studienpool A führte im Vergleich zum Abschlussbericht S07-01 jetzt dazu, dass sich

- für den Endpunkt Schulterdystokie ein **Beleg** für einen Nutzen auf der Basis einer statistisch signifikanten Meta-Analyse der Studien Crowther 2005 und Landon 2009 und
- für den Endpunkt Präeklampsie ein **Hinweis** auf einen Nutzen auf Basis eines statistisch signifikanten Unterschieds in Landon 2009 fand.

Zu den Endpunkten Kaiserschnitt, Geburtstraumata und neonatale Intensivpflege führte der Einschluss von Landon 2009 in die Meta-Analysen zu einer Verkleinerung der 95 %-Konfidenzintervalle, allerdings wurde keine statistische Signifikanz erreicht. Insbesondere bei dem seltenen Endpunkt Geburtstraumata könnte das ein Problem zu geringer statistischer Power sein.

Auch beim Surrogatparameter SGA blieb die Meta-Analyse nach Einschluss von Landon 2009 nicht statistisch signifikant, das Odds Ratio verschob sich jedoch von 1,05 auf 1,10. Während in Crowther 2005 bei Schwangeren mit ausgeprägter Glukosetoleranzstörung eine numerische Reduktion des Anteils der als SGA klassifizierten Kinder berichtet wurde, berichtete Landon 2009 bei Schwangeren mit offenbar milderer Glukosetoleranzstörung über eine numerische Erhöhung des Anteils der als SGA klassifizierten Kinder (Tabelle 15). In Landon 2009 wurden in der Interventionsgruppe etwa 8 % (37 / 477) der Frauen mit Insulin behandelt (BG-Zielwerte: nüchtern < 95 mg/dl; postprandial < 120 mg/dl). In Crowther 2005 waren es 20 % (100 / 490) bei weniger strengen Zielwerten (nüchtern \leq 99 mg/dl, postprandial \leq 126 mg/dl). Auch wenn das Risiko für SGA nach den Ergebnissen dieses Arbeitspapiers nicht stark erhöht zu sein scheint, fügen sich die Zahlen in das Bild, dass bei mildem Gestationsdiabetes ein Risiko einer Übertherapie besteht, wenn die Intensität der Therapie nicht dem Ausmaß der Störung der Glukosetoleranz angepasst wird.

Die Vermutung, dass eine GDM-spezifische Therapie zu einer Reduktion der perinatalen Mortalität führt, wird in Landon 2009 bei Frauen mit milderem GDM nicht bestätigt. Der Versuch einer meta-analytischen Zusammenfassung ergab einen deutlichen Hinweis auf Heterogenität. Die Ergebnisse aus Landon 2009 sind ein Argument, die in Crowther 2005 beobachtete Reduktion der perinatalen Mortalität als nicht repliziert und möglicherweise als Zufallsbefund zu bewerten. Hingegen wurde die in Crowther 2005 gesehene numerische Reduktion von Schulterdystokien und Geburtsverletzungen durch Landon 2009 bestätigt.

6 Diskussion

Durch den Einschluss von Landon 2009 bestätigt sich die Schlussfolgerung des Abschlussberichts S07-01, dass eine GDM-spezifische Therapie einen Nutzen haben kann. Für Schulterdystokien ergab sich ein Beleg für einen Nutzen. Wie bereits im Abschlussbericht aufgeführt, wird die Schulterdystokie als eine kritische geburtshilfliche Situation betrachtet, in der rasches Eingreifen erforderlich ist. Sie führt aber nur in einer Minderzahl der Fälle zu weiteren Komplikationen wie Geburtstraumata. So wurden in Crowther 2005 bei 22 von 23 Schulterdystokien keine weiteren Komplikationen berichtet, bei einem Kind kam es zu einer Nervenlähmung. In der Studie HAPO 2008 kam es bei 18,9 % der Kinder mit Schulterdystokie auch zu einem Geburtstrauma (siehe Abschlussbericht S07-01, Tabelle 57, Seite 162). HAPO 2008 zeigt aber auch, dass Schulterdystokien und kindliche Geburtstraumata selten sind. Sie traten im Durchschnitt bei 1,3 % der Geburten auf. Die Spannweite zwischen den Zentren reichte von 0,1 bis 3,4 %.

Landon 2009 hat insgesamt ein niedriges Verzerrungspotenzial; die Hauptschwäche ist, dass Entbindungsdaten von 8 Frauen in der Interventions- und 18 Frauen in der Kontrollgruppe vollständig fehlen. Für seltene Endpunkte wie perinatale Mortalität und kindliche Geburtstraumata führte das Fehlen dieser Daten dazu, dass das Verzerrungspotenzial als hoch eingeschätzt wurde.

Als Grund für das Fehlen der Daten gibt Landon 2009 an, dass die Schwangeren außerhalb der Studienzentren entbunden hätten. Genauere Gründe konnten die Autoren auf Anfrage nicht nennen, bis auf die nicht quantifizierte Aussage, dass einige Schwangere verzogen seien oder den Krankenversicherer gewechselt hätten. Es ist vorstellbar, dass auch akute Schwangerschaftskomplikationen ein Grund gewesen sein könnten, warum die Frauen nicht wie geplant in den Studienzentren entbunden haben. Dann wäre damit zu rechnen, dass auch Geburtskomplikationen häufiger bei den Schwangeren aufgetreten sind, deren Daten nicht erfasst wurden.

Es ist jedoch nicht davon auszugehen, dass durch das Fehlen solcher Daten die Ergebnisse der Studie zugunsten der Intervention verzerrt wurden. Da in der Kontrollgruppe von mehr Frauen (n = 18) Entbindungsdaten fehlen als in der Interventionsgruppe (n = 8), ist in diesem Szenario eher eine Unterschätzung zu erwarten.

Insgesamt wurde Landon 2009 deshalb als belastbare Quelle für das Teilziel 2 eingeschätzt.

Weiterhin keine direkten Aussagen zu Nutzen und Schaden eines Screenings möglich

Die Aktualisierungsrecherche hat zu Teilziel 1 „Screening“ keine Studien identifiziert. Auch durch den Einschluss von Landon 2009 ändert sich deshalb nichts an dem grundsätzlichen Dilemma, dass ohne verlässliche Screening-Studien keine direkten Aussagen zur Abwägung von Nutzen und Schaden eines Screenings auf Gestationsdiabetes möglich sind.

Insbesondere die Wissenslücken zu potenziellen Schäden eines Screenings, wie sie im Abschlussbericht S07-01 diskutiert wurden, werden durch den Einschluss von Landon 2009 nicht verringert. Allerdings ergeben sich auch keine neuen Argumente, die ein größeres Schadenspotenzial durch ein Screening auf Gestationsdiabetes erwarten lassen, als es im Abschlussbericht dargelegt wurde.

Insgesamt bleibt es bei der Einschätzung, dass indirekt ein Hinweis darauf abgeleitet werden kann, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt.

Übertragbarkeit

Auch nach Einschluss von Landon 2009 kann die Frage der Übertragbarkeit der Studienergebnisse nicht klarer beantwortet werden als im Abschlussbericht S07-01. Die in Landon 2009 eingeschlossenen Frauen waren im Durchschnitt etwa 29 Jahre alt, 20 bis 25 % haben zum ersten Mal entbunden, der BMI lag im Mittel bei etwa 30 kg/m². Etwa 57 % der Frauen waren spanischstämmig, ein Viertel war „weiß“ (Kaukasierinnen). In Ethnie und BMI unterscheidet sich die Studienpopulation damit deutlich von deutschen Schwangeren [26,27]. In Deutschland wiesen 2008 11,4 % der Schwangeren einen BMI über 30 kg/m² auf, bei 82 % ist das Herkunftsland Deutschland [26].

Welche Bedeutung die Eigenschaften der Population für den Zusammenhang zwischen Blutglukosekonzentration und Risiken für Mutter und Kind haben, zeigt sich an einem Vergleich der Nüchtern-BG-Werte der Studien Crowther 2005 und Landon 2009. Trotz der unterschiedlichen Auswahlkriterien war die mittlere Nüchtern-BG der Kontrollgruppe in Landon 2009 mit 86,3 mg/dl (SD 5,7 mg/dl) praktisch gleich der mittleren Nüchtern-BG der Kontrollgruppe in Crowther 2005 (86,5 mg/dl; SD 10,8 mg/dl). Trotzdem unterscheidet sich der LGA-Anteil zwischen den Studien erheblich: 14,5 % in Landon 2009 vs. 22 % in Crowther 2005. Das bestätigt den bereits aus dem Vergleich von Crowther 2005 mit HAPO 2008 gezogenen Schluss, dass die prognostische Bedeutung der Blutglukosewerte je nach weiteren Charakteristika sehr unterschiedlich ausfallen kann.

Aus diesen Gründen ist mit den aus den Publikationen zur Verfügung stehenden Daten kein belastbarer Abgleich der Ergebnisse von Landon 2009 und HAPO 2008 [28] möglich.

Selektion der Teilnehmerinnen in Landon 2009

Eine Gemeinsamkeit aller Studien in Studienpool A ist, dass die Selektion der Frauen in 2 Stufen stattfand, insbesondere durch einen GCT als Vortest. In Crowther 2005 konnten Frauen zwar auch aufgrund der Abfrage von Risikofaktoren für einen oGTT infrage kommen, allerdings wurden auf diese Weise nur 7 % der Teilnehmerinnen in die Studie eingeschlossen, 93 % wiesen einen positiven 1-Stunden-GCT auf. In Bonomo 2005 und Crowther 2005 lag

der Grenzwert für einen positiven GCT bei ≥ 140 mg/dl, in Langer 1989 und O'Sullivan 1966 bei ≥ 130 mg/dl, in Landon 2009 bei ≥ 135 mg/dl.

In Landon 2009 werden auch Einzelheiten der Selektion der Teilnehmerinnen berichtet, aus denen sich Abschätzungen des gesamten Selektionsprozesses ableiten lassen.

Insgesamt wurden in Landon 2009 19 665 Frauen mit einem GCT ≥ 135 mg/dl identifiziert. Von diesen 19 665 Schwangeren wurden folgende große Gruppen ausgeschlossen und absolvierten keinen oGTT;

- 8676 (44,1 %) erfüllten nicht die weiteren Einschlusskriterien (siehe Tabelle 8).
- 3554 (18,1 %) lehnten eine Teilnahme ab.

Insgesamt absolvierten 7298 Frauen einen 3-Stunden-oGTT. Davon hatten

- 4348 einen normalen oGTT (GCT falsch-positiv),
- 1938 eine NüBG > 95 mg/dl und erhielten eine Therapie,
- 54 nicht die Einschlusskriterien der Studie erfüllt und
- 958 einen „abnormalen“ oGTT (GCT richtig-positiv). Diese Frauen wurden randomisiert.

Toronto 1998 berichtete, dass die Spezifität des Trennwertes 135 mg/dl in einem 1-Stunden-GCT bei 76,7 % lag (Abschlussbericht S07-01, Tabelle 97, S. 240). Unter Berücksichtigung von Prävalenz und Sensitivität wies also etwa 1 von 4 Frauen einen positiven GCT auf.

Daraus lässt sich grob abschätzen, dass für die Selektion der Teilnehmerinnen an der Studie Landon 2009 etwa 80 000 Frauen einen 1-Stunden-GCT absolvierten ($\sim 4 \times 19\ 665$).

Nach dieser Abschätzung hätten von 80 000 Schwangeren:

- etwa 20 000 (25 %) einen positiven GCT
- etwa 1938 (2,4 %) einen positiven GCT **und** einen NüBG > 95 mg/dl,
- etwa 958 (1,2 %) einen positiven GCT **und** einen positiven 3-Stunden-100-g-oGTT (nach Carpenter-Coustan- / ADA-Kriterien).

Bei Frauen mit einem positiven GCT **und** einem NüBG > 95 mg/dl wurde in Landon 2009 ebenfalls GDM diagnostiziert und behandelt, über die Ergebnisse berichtet Landon 2009 jedoch nicht.

Insgesamt würde also nach dieser Schätzung die von Landon 2009 gewählte Strategie bei 3,6 % der Schwangeren zu einer GDM-Therapie führen. Es ist jedoch zu erwarten, dass auch

in den Gruppen der Frauen, die die Einschlusskriterien nicht erfüllten oder die Teilnahme ablehnten, weitere Schwangere die Kriterien für eine GDM-Diagnose erfüllt hätten. Demnach läge der Anteil der Frauen mit einer GDM-Diagnose also höher.

Zu einer deutlich anderen Abschätzung kommt man, wenn man die Ergebnisse des Surrogatparameters LGA zu einem Vergleich der Populationen von Landon 2009 und HAPO 2008 heranzieht. Der Anteil der als LGA klassifizierten Kinder betrug in der Kontrollgruppe von Landon 2009 14,5 %. Ein LGA-Anteil in vergleichbarer Höhe wie in Landon 2009 wurde in HAPO 2008 bei Frauen ab Kategorie 4 der Nüchtern-BG (> 85 mg/dl) erreicht. Der Anteil der als LGA klassifizierten Kinder lag in den Kategorien 4 bis 7 bei 14,8 % (814 LGA-Kinder auf 5513 Geburten) (siehe Abschlussbericht S07-01, Tabelle 53, S. 158 und Tabelle 64, S. 168). In diesen Kategorien befanden sich 23,7 % der HAPO-Population.

Die Tatsache, dass diese beiden Abschätzungen sehr unterschiedliche Ergebnisse liefern, ist ein weiteres Indiz dafür, dass die Verwendung des 1-Stunden-GCT als Vortest eine Population selektiert, auf die die Risikozusammenhänge aus HAPO 2008 nicht ohne weiteres übertragen werden können.

Aspekte der Ableitung einer Screeningstrategie

Die im Folgenden dargestellten Überlegungen sollen mehr der Beschreibung der Unsicherheiten dienen als der Ableitung einer plausiblen Screeningstrategie. Auf Basis der Publikationen von Crowther 2005, Landon 2009 und HAPO 2008 lässt sich keine Screeningstrategie eindeutig als optimal ableiten. Es gelten weiterhin die im Abschlussbericht S07-01 ab Seite 235 beschriebenen Erwägungen.

Mit Landon 2009 zeichnen sich weiterhin alle Therapiestudien des Pools A dadurch aus, dass die Vorauswahl fast ausschließlich durch einen 1-Stunden-50-g-GCT als Vortest stattfand. Die Ergebnisse von Landon 2009 sprechen dafür, für diesen Test eine Schwelle von 135 mg/dl anzusetzen, da schon für die niedrigere Schwelle ein Effekt gezeigt ist.

Wie viele schwangere Frauen in Deutschland diese Schwelle überschreiten würden, ist unklar. Auf der Basis von Toronto 1998 lässt sich abschätzen, dass etwa 25 bis 30 % der Schwangeren über dieser Schwelle liegen würden (siehe Tabellen 97 und 98, S. 240 des Abschlussberichts S07-01).

Bühling [29] berichtet über den Einsatz eines 1-Stunden-50-g-GCT im Zeitraum September 1994 bis Juni 1996 bei einer Gruppe von 1416 Schwangeren, die die Schwangerenberatung der Universitätsfrauenklinik der Berliner Charité aufsuchten (ca. 69 % Westeuropäerinnen, ca. 26 % Türkinen). Hier wurde bei 26,3 % der Schwangeren der Grenzwert von 140 mg/dl überschritten.

Niedriger liegt der Anteil in einer niederländischen Studie [30]. In der Studie Van Leeuwen et al. 2007 absolvierten 1301 Schwangere in der 24. bis 28. SSW einen 50-g-GCT, der 1-Stunden-Grenzwert lag bei > 140 mg/dl. In Van Leeuwen et al. 2007 lagen 167 (13,0 %) über diesem Grenzwert. Von diesen 167 Frauen hatten wiederum 33 (19,8 %) einen auffälligen 75-g-oGTT nach WHO-Kriterien. Letztlich erhielten durch diese Strategie 2,5 % der Frauen (33 von 1301) eine GDM-Diagnose.

Möglicherweise lassen sich auch weitere deutsche Daten (zum Beispiel von Bühling oder Kleinwechter [31,32]) dazu nutzen, für Deutschland genauere Schätzungen der Testcharakteristika des Schwellenwerts von 135 mg/dl in einem 1-Stunden-GCT abzuleiten.

Sollte ein Screeningprogramm in Deutschland eingeführt werden, sollte sich auch die Definition der Grenzwerte für den auf den GCT folgenden oGTT an den Ergebnissen des Studienpools A orientieren. Die Ableitung von diagnostischen Kriterien aus Landon 2009 wird jedoch dadurch erschwert, dass der in Landon 2009 verwendete 100-g-oGTT in Deutschland nicht üblich ist. Es existiert keine validierte „Formel“ dafür, die Grenzwerte eines 100-g-oGTT in die für einen 75-g-oGTT zu überführen.

Mehrere Gruppen haben Vergleiche der BG-Messwerte nach Belastung mit 100 g oder 75 g Glukose publiziert [33-35]. Das Konzept dieser Studien war, dass Schwangere innerhalb von 1 bis 2 Wochen sowohl einen 75-g- als auch einen 100-g-oGTT absolviert haben. Ein übereinstimmendes Ergebnis ist, dass die höhere Glukosebelastung auch zu höheren BG-Werten nach 1, 2 und 3 Stunden führte. Allerdings fielen die Differenzen in den Studien unterschiedlich aus, sodass sich keine einfache Umrechnungsformel ableiten lässt. In Mello et al. [33] lagen die BG-Werte 1 Stunde nach einem 100-g-oGTT etwa 4 mg/dl und nach 2 Stunden etwa 7 mg/dl höher als nach einem 75-g-oGTT. In Weiss et al. [35] lagen die BG-Werte 1 und 2 Stunden nach einem 100-g-oGTT bis zu 16 mg/dl höher als nach einem 75-g-oGTT. In Brustman et al. [34] lagen die BG-Werte 1 Stunde nach einem 100-g-oGTT bis zu 13 mg/dl, nach 2 Stunden bis zu 22 mg/dl und nach 3 Stunden bis zu 23 mg/dl höher als nach einem 75-g-oGTT.

Physiologische Untergruppen

Ein zweiter Aspekt von Landon 2009 ist, dass Schwangere mit einer NüBG ≥ 95 mg/dl nicht an der Studie teilnahmen, weil die Autoren für diese Frauen eine Randomisierung und einen Verzicht auf eine Therapie nicht für angemessen hielten. Da Erhöhungen der Nüchternblutglukosewerte als Zeichen einer ausgeprägteren Glukosestoffwechselstörung gelten [36], ist diese Logik prinzipiell nachvollziehbar.

Weitere Einzelheiten sind allerdings für diese Gruppe von Frauen nicht berichtet. Dieser Ausschluss von Frauen aufgrund von NüBG-Werten führte jedoch möglicherweise dazu, dass in Landon 2009 Frauen eingeschlossen waren, die eher durch eine postprandiale Störung der Glukosetoleranz charakterisiert waren.

Diese Selektion von Frauen mit einer Störung der postprandialen Glukosetoleranz wird auch durch den GCT als Vortest noch verstärkt. Es ist beschrieben, dass der Test je nach Abstand zur letzten Mahlzeit variable Ergebnisse zeigt [29,37].

Diese Abhängigkeit der Ergebnisse des 50-g-GCT je nach Abstand zur letzten Mahlzeit führt jedoch nicht dazu, dass Frauen mit einer Störung des Glukosestoffwechsels übersehen werden. Bei Frauen, die unmittelbar vor dem 50-g-GCT eine kohlenhydratreiche Mahlzeit zu sich genommen haben, kann das auch bei normaler Glukosetoleranz zu einem auffälligen GCT-Ergebnis führen. Das wäre ein falsch-positives Ergebnis und würde durch einen normalen oGTT entkräftet.

Wenn Frauen den 50-g-GCT nüchtern absolvieren, macht es der niedrig liegende Trennwert des GCT unwahrscheinlich, dass Frauen übersehen würden, die in einem oGTT einer Störung der 1-Stunden-Glukosetoleranz aufweisen. Der in den Studien des Pools A verwendete Grenzwert des 50-g-GCT liegt mit maximal 140 mg/dl viel niedriger als der 1-Stunden-Wert für einen 75-g-oGTT (170 bis 190 mg/dl, je nach Definition; siehe Tabelle 3, S. 8 des Abschlussberichts S07-01).

Der GCT als Vortest führt aber dazu, dass Frauen nicht identifiziert werden, deren Glukosestoffwechselstörung durch eine isoliert erhöhte NüBG gekennzeichnet ist. Diese Selektion wurde in Crowther 2005 dadurch verstärkt, dass ein erhöhter 2-Stunden-Wert maßgebliches Einschlusskriterium war und in Landon 2009 eine Kombination von 2 erhöhten 1-, 2- oder 3-Stunden-Werten für den Einschluss in die Studie vorliegen musste.

Aus diesen Überlegungen lässt sich zusammenfassend ableiten, dass der in den Therapiestudien Crowther 2005 und Landon 2009 eingesetzte 50-g-Vortest und die Diagnosekriterien des oGTT dazu geführt haben, dass Frauen mit Glukosestoffwechselstörungen, die durch eine isoliert erhöhte NüBG gekennzeichnet sind, nicht an den Studien teilgenommen haben.

Es ist jedoch unklar, wie groß in den Studien der Anteil der Frauen ist, die eine isolierte Erhöhung der NüBG aufwiesen.

Aspekte des „HAPO-Konsensus“

Eine internationale Konsensusgruppe hat angekündigt, auf der Basis der Daten von HAPO 2008 einen Vorschlag für das Screening und die Diagnose des Gestationsdiabetes abzuleiten. Laut Ankündigung auf dem 5th International Symposium on Diabetes and Pregnancy Ende März 2009 in Sorrent, Italien [38] soll ein standardisierter 75-g-oGTT für alle Schwangeren in der 24. bis 28. Schwangerschaftswoche eingesetzt werden, mit den in Tabelle 23 beschriebenen Grenzwerten. Der Konsens beinhaltet, dass die Diagnose GDM gestellt wird, wenn mindestens **ein** Wert überschritten wird.

Bei einer Nüchtern-BG von 92 mg/dl in der 24. bis 28. SSW (venöses Plasma) würden 8,3 % der HAPO-Population die Diagnose GDM erhalten. Über diesem Wert oder einem 1-Stunden-Wert von 180 mg/dl lagen 13,9 % der HAPO-Population. Über einem dieser beiden Werte oder einem 2-Stunden-Wert von 153 mg/dl lagen 16,0 %.

Bei 604 der 3726 Frauen (16,2 %) war das Kind als LGA klassifiziert. Auf diese Weise würden etwa 27 % aller in HAPO 2008 als LGA klassifizierten Kinder erfasst.

Tabelle 23: HAPO 2008 – Konsensusvorschlag für Grenzwerte eines Screenings auf Gestationsdiabetes

Messzeitpunkt	Grenzwerte (mg/dl)	Zahl der Frauen „GDM“ kumuliert (%)	Zahl der Kinder „LGA“ kumuliert (%)
Nüchtern	92	1925 (8,3)	376 (16,9)
1 h	180	3238 (13,9)	535 (24,1)
2 h	153	3726 (16,0)	604 (27,2)
Gesamtzahl		23 316	2221

Dieser Vorschlag hätte zur Folge, dass

1. kein Vortest stattfände.
2. Schwangere alleine auf Basis einer erhöhten NüBG als GDM eingestuft werden könnten. Für diese Frauen liegen keine Daten aus Therapiestudien vor.
3. Blutglukose-Grenzwerte für einen 75-g-oGTT eingesetzt würden, die fast denen entsprechen, die in Landon 2009 für einen 100-g-oGTT verwendet wurden. Das würde Frauen mit einer eher stärker ausgeprägten Störung des Glukosestoffwechsels identifizieren. Allerdings müsste im Unterschied zu Landon 2009 nur ein Wert überschritten sein. Das würde eher Frauen mit einer geringer ausgeprägten Störung des Glukosestoffwechsels identifizieren, die in Landon 2009 ausgeschlossen waren.
4. für den 2-Stunden-Wert eine höhere Schwelle definiert würde als in Crowther 2005 (153 statt 140 mg/dl). Das würde Frauen mit einer eher stärker ausgeprägten Störung des Glukosestoffwechsels identifizieren.

Die Unterschiede sind in ihrer Wirkung zum Teil gegenläufig, sodass keine Vorhersage möglich ist, inwieweit eine durch diese Konsenskriterien ausgewählte Population den Teilnehmerinnen in den Therapiestudien gleichen würde.

Für den HAPO-Konsensus herangezogene Endpunkte

Das Vorgehen zur Festlegung der Grenzwerte im HAPO-Konsens ist einer Arbeit von Coustan et al. skizziert [39]. Die Autoren beschreiben, dass für die Ableitung der diagnostischen Grenzwerte die Assoziationen zwischen den BG-Werten zum Zeitpunkt nüchtern, 1 Stunde und 2 Stunden und den Häufigkeiten von „Large-for-Gestational-Age“, „Prozent Körperfett > 90. Perzentile“ und „Hyperinsulinämie (Nabelschnurblut C-Peptid > 90. Perzentile)“ festgelegt wurden.

Keiner dieser Endpunkte wurde für das Projekt S07-01 als patientenrelevant eingestuft. Die diagnostischen Grenzwerte des Konsensus beruhen auf Surrogatparametern fraglicher klinischer Relevanz.

Im zweiten Schritt des Konsensprozesses wurde festgelegt, die mittleren Blutglukosewerte zu den Zeitpunkten Nüchtern, 1 Stunde und 2 Stunden als Ausgangspunkt zu wählen. Als „behandlungsbedürftiger GDM“ wurden die Blutglukosewerte definiert, bei denen das Risiko für die 3 oben beschriebenen Risiken um ein OR von 1,75 erhöht war.

Tabelle 24 enthält die in HAPO 2008 berichteten Assoziationen zwischen der Höhe der Blutglukose und den Risiken für ausgewählte Endpunkte. Auffällig ist, dass sich zu allen Zeiten die höchsten OR für die Endpunkte Hyperinsulinämie (C-Peptid > 90. Perzentile) und LGA fanden. Die Ableitung von BG-Grenzwerten, bei denen ein OR von 1,75 erreicht ist, führt auf Basis dieser Endpunkte deshalb zu niedrigeren BG-Grenzwerten, als es für andere Endpunkte, zum Beispiel Geburtstraumata der Fall wäre.

Tabelle 24: HAPO 2008 – Assoziationen zwischen BG und Risiken für ausgewählte Endpunkte

Endpunkt	NüBG adjustiertes OR^a	1 Stunde adjustiertes OR^a	2 Stunden adjustiertes OR^a
LGA	1,38 (1,32;1,44)	1,46 (1,39;1,53)	1,38 (1,32;1,44)
Erstkaiserschnitt ^b	1,11 (1,06;1,15)	1,10 (1,06;1,15)	1,08 (1,03;1,12)
klinische neonatale Hypoglykämie	1,08 (0,98;1,19) ^c	1,13 (1,03;1,26)	1,10 (1,00;1,12)
Nabelschnur-C-Peptide > 90. Perzentile	1,55 (1,47;1,64)	1,46 (1,38;1,54)	1,37 (1,30;1,44)
Frühgeburt (vor 37. SSW)	1,05 (0,99;1,11)	1,18 (1,12;1,25)	1,16 (1,10;1,23)
Schulterdystokie oder Geburtstrauma	1,18 (1,04;1,33)	1,23 (1,09;1,38)	1,22 (1,09;1,37)
neonatale Intensivpflege	0,99 (0,94;1,05)	1,07 (1,02;1,13)	1,09 (1,03;1,14)
Hyperbilirubinämie	1,00 (0,95;1,05)	1,11 (1,05;1,17)	1,08 (1,02;1,13)
Präeklampsie	1,21 (1,13;1,29)	1,28 (1,20;1,37)	1,28 (1,20;1,37)
a: Diese Odds Ratios gelten für einen Anstieg der Plasmablutglukose um 1 SD (6,9 mg/dl nüchtern, 30,9 mg/dl 1 h und 23,5 mg/dl 2 h). b: Daten der Frauen mit früheren Kaiserschnitten wurden ausgeschlossen. c: Der p-Wert für die quadratische (nicht lineare) Assoziation war 0,013.			

Nach den Ergebnissen dieses Arbeitspapiers sind diese 3 Endpunkte nicht gleich gut durch eine Therapie zu beeinflussen. Für LGA findet sich wie in dem vorliegenden Arbeitspapier beschrieben, ein deutlicher Effekt in den Studien des Pools A. Landon 2009 ist die einzige Studie, die zur Fettmasse (als kontinuierliche Größe) und zu Hyperinsulinämie berichtet. Die mittlere Fettmasse lag in Landon 2009 in der Interventionsgruppe bei 427,0 g (SD 197,9) und in der Kontrollgruppe bei 464,3 g (SD 222,3). Der Unterschied der Gruppenmittel von 37,3 g war statistisch signifikant (p = 0,003).

Der Endpunkt Hyperinsulinämie (C-Peptid > 95. Perzentile) war in Landon 2009 Teil des primären kombinierten Endpunkts. Die Rate war in der Interventionsgruppe nicht statistisch signifikant verringert.

Damit fehlt für 1 der 3 für die Festlegung des Konsenses gewählten Endpunkte ein Beleg dafür, dass er durch eine Therapie zu beeinflussen ist.

Zusammenfassend hat die Ableitung des HAPO-Konsensus folgende methodische Aspekte:

- Es wurden zur Festlegung der Risikoschwelle keine patientenrelevanten Endpunkte gewählt.
- Für den Surrogatendpunkt Hyperinsulinämie (C-Peptid > 95. Perzentile) ist unbewiesen, dass er durch eine Therapie klinisch bedeutsam beeinflusst werden kann.
- Für den Endpunkt Fettmasse ist unklar, ob die Reduktion eine klinische Bedeutung hat.

Diese Aspekte führen dazu, dass die Grenzwerte im HAPO-Konsensus niedriger angesetzt wurden, als es bei der Orientierung an patientenrelevanten Endpunkten der Fall gewesen wäre.

Unterschiede zwischen Therapiestudien und HAPO-Konsens

An den maßgeblichen Therapiestudien haben Frauen teilgenommen, deren GDM-Diagnose aufgrund der Kombination von GCT-Vortest und Diagnosekriterien im oGTT vor allem auf einer postprandialen Glukosestoffwechselstörung beruht.

Einerseits ist zu erwarten, dass der HAPO-Konsens einen Teil der Frauen nicht erfasst, der in den Therapiestudien eingeschlossen wurde.

- Im Vergleich zu Crowther 2005 liegt der 2-Stunden-Wert höher (153 statt 140 mg/dl).
- Im Vergleich zu Landon 2009 wird ein 75-g-oGTT statt eines 100-g-oGTT eingesetzt, dadurch sind gerade bei Frauen mit einer Störung des Glukosestoffwechsels niedrigere Blutzuckerwerte zu erwarten.

Andererseits ist zu erwarten, dass der HAPO-Konsens Frauen erfasst, die in den Therapiestudien nicht eingeschlossen wurden.

- Im Vergleich zu Crowther 2005 und Landon 2009 werden Frauen erfasst, die einen unauffälligen GCT aufweisen.
- Im Vergleich zu Crowther 2005 und Landon 2009 werden Frauen alleine aufgrund erhöhter Nüchtern-BG im oGTT erfasst.

Es sind in diesem Arbeitspapier keine Abschätzungen möglich, wie groß die Überlappungen sind zwischen der Population, die durch den HAPO-Konsens ausgewählt wurde, und den Populationen, die in den maßgeblichen Therapiestudien teilgenommen haben.

Forschungsbedarf

Der Abschlussbericht S07-01 beschreibt bereits den bestehenden Forschungsbedarf. An der Situation ändert sich durch den Einschluss von Landon 2009 in Studienpool A grundsätzlich nichts. Es fehlen weiterhin wichtige Informationen, um zu entscheiden, welche Screeningstrategie auf Gestationsdiabetes für Schwangere in Deutschland die beste Abwägung von möglichem Nutzen und möglichem Schaden bietet.

Im Rahmen zukünftiger Forschungsprojekte könnte auch die unten skizzierte zweistufige Strategie erprobt werden. Basis wäre die Risikoeinschätzung im Rahmen der Mutterschaftsrichtlinien, bei der sich keine besonderen Risiken ergeben haben:

1. Stufe: ab 24. bis 28. SSW: Vortest 50-g-GCT:

wenn 1-Stunden-Wert ≥ 135 mg/dl

→ **2. Stufe:** 24. bis 30. SSW 75-g-oGTT:

wenn 1 Wert erhöht: NüBG ≥ 95 mg/dl ODER 2-Stunden-BG ≥ 140 mg/dl

→ → **Diagnose GDM**

Diese Strategie übernimmt von Landon 2009 den 1-Stunden-50-g-GCT-Grenzwert ≥ 135 mg/dl und den NüBG-Grenzwert ≥ 95 mg/dl. Sie übernimmt von Crowther 2005 den 2-Stunden-75-g-oGTT-Grenzwert ≥ 140 mg/dl. In Landon 2009 lag der 2-Stunden-100-g-oGTT-Grenzwert bei ≥ 155 mg/dl. Der 2-Stunden-Grenzwert aus Crowther 2005 würde also eine Anpassung an die niedrigere Glukosebelastung beinhalten und wäre auch mit Landon 2009 vereinbar.

Für einen 1-Stunden-Grenzwert ist keine Ableitung aus den Therapiestudien möglich, da er in Crowther 2005 keine Rolle spielte und der 1-Stunden-Grenzwert ≥ 180 mg/dl in Landon 2009 auf einem 100-g-oGTT beruhte. Der HAPO-Vorschlag ≥ 180 mg/dl 1 Stunde nach einem 75-g-oGTT würde also Frauen mit einer stärker ausgeprägten Glukosestoffwechselstörung als in Landon 2009 selektieren.

7 Abschätzung der Auswirkungen auf das Fazit des Abschlussberichtes S07-01

Das Fazit des Abschlussberichtes S07-01 lautete:

Es liegt ein Hinweis auf einen Nutzen einer Gestationsdiabetes-spezifischen Therapie vor.

Es existieren keine direkten Belege für oder Hinweise auf einen Nutzen oder einen Schaden durch ein Screening auf Gestationsdiabetes, da keine geeigneten Screeningstudien identifiziert wurden.

Trotzdem kann indirekt ein Hinweis darauf abgeleitet werden, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt.

Die Ergebnisse dieses Arbeitspapiers würden zu folgender Änderung des Fazits des Abschlussberichtes S07-01 führen (Änderungen kursiv):

Es liegt ein *Beleg* für den Nutzen einer Gestationsdiabetes-spezifischen Therapie vor.

Es existieren keine direkten Belege für oder Hinweise auf einen Nutzen oder einen Schaden durch ein Screening auf Gestationsdiabetes, da keine geeigneten Screeningstudien identifiziert wurden.

Trotzdem kann indirekt ein Hinweis darauf abgeleitet werden, dass ein Screening auf Gestationsdiabetes zu einer Reduktion von perinatalen Komplikationen führt.

8 Liste der eingeschlossenen Studien

8.1 Liste der durch die Nachrecherche eingeschlossenen Studien

Teilziel 1 „Screening“

keine

Teilziel 2 „Therapie“, Studienpool A

Landon 2009

Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B et al. A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med* 2009; 361(14): 1339-1348.

Landon MB, Thom E, Spong CY, Carpenter M, Mele L, Johnson F et al. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Unit Network randomized clinical trial in progress: standard therapy versus no therapy for mild gestational diabetes. *Diabetes Care* 2007; 30(Suppl 2): S194-S199.

Landon MB, Thom E, Spong CY, Gabbe SG, Leindecker S, Johnson F et al. A planned randomized clinical trial of treatment for mild gestational diabetes mellitus. *J Matern Fetal Neonatal Med* 2002; 11(4): 226-231.

8.2 Liste der bereits im Abschlussbericht S07-01 eingeschlossenen Studien zum Teilziel 2 „Therapie“, Studienpool A

Bonomo 2005

Bonomo M, Corica D, Mion E, Goncalves D, Motta G, Merati R et al. Evaluating the therapeutic approach in pregnancies complicated by borderline glucose intolerance: a randomized clinical trial. *Diabet Med* 2005; 22(11): 1536-1541.

Coustan 1978

Coustan DR, Lewis SB. Insulin therapy for gestational diabetes. *Obstet Gynecol* 1978; 51(3): 306-310.

Crowther 2005

Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. *N Engl J Med* 2005; 352(24): 2477-2486.

Athukorala C, Crowther CA, Willson K. Women with gestational diabetes mellitus in the ACHOIS trial: risk factors for shoulder dystocia. *Aust N Z J Obstet Gynaecol* 2007; 47(1): 37-41.

Moss JR, Crowther CA, Hiller JE, Willson KJ, Robinson JS. Costs and consequences of treatment for mild gestational diabetes mellitus: evaluation from the ACHOIS randomised trial. *BMC Pregnancy Childbirth* 2007; 7: 27.

Langer 1989

Langer O, Levy J, Brustman L, Anyaegbunam A, Merkatz R, Divon M. Glycemic control in gestational diabetes mellitus: how tight is tight enough; small for gestational age versus large-for-gestational-age? *Am J Obstet Gynecol* 1989; 161(3): 646-653.

Li 1987

Li DFH, Wong VCW, O'Hoy KMKY. Is treatment needed for mild impairment of glucose tolerance in pregnancy? A randomized controlled trial. *Br J Obstet Gynaecol* 1987; 94(9): 851-854.

Corcoy R, Codina M, De Leiva A, Cerqueira MJ, Cabero L. Is treatment needed for mild impairment of glucose tolerance in pregnancy? A randomized controlled trial. *Br J Obstet Gynaecol* 1988; 95(4): 425-426.

Fraser R. Is treatment needed for mild impairment of glucose tolerance in pregnancy? A randomized controlled trial. *Br J Obstet Gynaecol* 1988; 95(6): 636-637.

O'Sullivan 1966

O'Sullivan JB, Gellis SS, Dandrow RV, Tenney BO. The potential diabetic and her treatment in pregnancy. *Obstet Gynecol* 1966; 27: 683-689.

O'Sullivan JB. Prospective study of gestational diabetes and its treatment. In: Sutherland HW, Stowers JM (Ed). *Carbohydrate metabolism in pregnancy and the newborn*. Edinburgh: Churchill Livingstone; 1975. S. 195-204.

O'Sullivan JB, Mahan CM. Insulin treatment and high risk groups. *Diabetes Care* 1980; 3(3): 482-485.

9 Literatur

1. Arbeitsgemeinschaft Diabetes und Schwangerschaft der Deutschen Diabetesgesellschaft, Arbeitsgemeinschaft für materno-fetale Medizin (AGMFM) der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe (DGGG), Deutsche Gesellschaft für Perinatale Medizin. Empfehlungen zu Diagnostik und Therapie des Gestationsdiabetes (GDM) [Online]. 2001 [Zugriff am 09.08.2007]. URL: <http://www.uni-duesseldorf.de/AWMF/II/057-008.htm>.
2. Metzger BE, Coustan DR. Summary and recommendations of the Fourth International Workshop-Conference on Gestational Diabetes Mellitus. *Diabetes Care* 1998; 21(Suppl 2): B161-B167.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Screening auf Gestationsdiabetes: Abschlussbericht; Auftrag S07-01; Version 1.1 [Online]. 25.08.2009 [Zugriff am 10.02.2010]. URL: http://www.iqwig.de/download/S07-01_Abschlussbericht_Screening_auf_Gestationsdiabetes.pdf.
4. Landon MB, Thom E, Spong CY, Carpenter M, Mele L, Johnson F et al. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Unit Network randomized clinical trial in progress: standard therapy versus no therapy for mild gestational diabetes. *Diabetes Care* 2007; 30(Suppl 2): S194-S199.
5. Landon MB, Thom E, Spong CY, Gabbe SG, Leindecker S, Johnson F et al. A planned randomized clinical trial of treatment for mild gestational diabetes mellitus. *J Matern Fetal Neonatal Med* 2002; 11(4): 226-231.
6. Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B et al. A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med* 2009; 361(14): 1339-1348.
7. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials 2001. *Explore (NY)* 2005; 1(1): 40-45.
8. Des Jarlais D, 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.
9. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003; 49(1): 1-6.

10. Vandembroucke JP, Von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007; 4(10): e297.
11. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandembroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007; 4(10): e296.
12. Bonomo M, Corica D, Mion E, Goncalves D, Motta G, Merati R et al. Evaluating the therapeutic approach in pregnancies complicated by borderline glucose intolerance: a randomized clinical trial. *Diabet Med* 2005; 22(11): 1536-1541.
13. Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. *N Engl J Med* 2005; 352(24): 2477-2486.
14. Athukorala C, Crowther CA, Willson K. Women with gestational diabetes mellitus in the ACHOIS trial: risk factors for shoulder dystocia. *Aust N Z J Obstet Gynaecol* 2007; 47(1): 37-41.
15. Moss JR, Crowther CA, Hiller JE, Willson KJ, Robinson JS. Costs and consequences of treatment for mild gestational diabetes mellitus: evaluation from the ACHOIS randomised trial. *BMC Pregnancy Childbirth* 2007; 7: 27.
16. Langer O, Levy J, Brustman L, Anyaegbunam A, Merkatz R, Divon M. Glycemic control in gestational diabetes mellitus: how tight is tight enough; small for gestational age versus large for gestational age? *Am J Obstet Gynecol* 1989; 161(3): 646-653.
17. Herbison P, Wilson D. Intensified versus conventional management of gestational diabetes. *Am J Obstet Gynecol* 1995; 172(5): 1642-1643.
18. O'Sullivan JB, Gellis SS, Dandrow RV, Tenney BO. The potential diabetic and her treatment in pregnancy. *Obstet Gynecol* 1966; 27: 683-689.
19. O'Sullivan JB. Prospective study of gestational diabetes and its treatment. In: Sutherland HW, Stowers JM (Ed). *Carbohydrate metabolism in pregnancy and the newborn*. Edinburgh: Churchill Livingstone; 1975. S. 195-204.
20. O'Sullivan JB, Mahan CM. Insulin treatment and high risk groups. *Diabetes Care* 1980; 3(3): 482-485.
21. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979; 28(12): 1039-1057.

22. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol* 1992; 31(Pt 3): 301-306.
23. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987; 150: 782-786.
24. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36); I: conceptual framework and item selection. *Med Care* 1992; 30(6): 473-483.
25. Andres AM, Mato AS. Choosing the optimal unconditioned test for comparing two independent proportions. *Comput Stat Data Anal* 1994; 17(5): 555-574.
26. Bundesgeschäftsstelle Qualitätssicherung. BQS-Bundesauswertung 2008: Geburtshilfe [Online]. 2009 [Zugriff am 12.01.2010]. URL: http://www.bqs-outcome.de/2008/ergebnisse/leistungsbereiche/geburtshilfe/buaw/bu_Gesamt_16n1-GEBH_2008.pdf.
27. Bauer S, Fischer B, Sandu C, Vetter K. Mütterliches und kindliches Outcome bei Gestationsdiabetes. *Frauenarzt* 2009; 50(2): 130-139.
28. Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, Coustan DR et al. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med* 2008; 358(19): 1991-2002.
29. Bühling KJ, Stein U, Dudenhausen JW. Evaluation des 50 g-Glukose-Screeningtests an 1416 Schwangeren. *Geburtshilfe Frauenheilkd* 1998; 58(2): 100-109.
30. Van Leeuwen M, Zweers EJ, Opmeer BC, Van Ballegoie E, Ter Brugge HG, De Valk HW et al. Comparison of accuracy measures of two screening tests for gestational diabetes mellitus. *Diabetes Care* 2007; 30(11): 2779-2784.
31. Bühling KJ, Dudenhausen JW. Erkennung des Gestationsdiabetes. *Zentralbl Gynakol* 2003; 125: 123-128.
32. Kleinwechter HJ. Pilotprojekt Gestationsdiabetes Schleswig-Holstein: Zwischenbilanz und Durchführbarkeit nach 16 Monaten. *Schleswig-Holsteinisches Ärzteblatt* 2000; (2): 21-24.
33. Mello G, Elena P, Ognibene A, Cioni R, Tondi F, Pezzati P et al. Lack of concordance between the 75-g and 100-g glucose load tests for the diagnosis of gestational diabetes mellitus. *Clin Chem* 2006; 52(9): 1679-1684.
34. Brustman LE, Gela BD, Moore M, Reilly KD, Langer O. Variations in oral glucose tolerance tests: the 100- versus 75-g controversy. *J Assoc Acad Minor Phys* 1995; 6(2): 70-72.

35. Weiss PA, Haeusler M, Kainer F, Purstner P, Haas J. Toward universal criteria for gestational diabetes: relationships between seventy-five and one hundred gram glucose loads and between capillary and venous glucose concentrations. *Am J Obstet Gynecol* 1998; 178(4): 830-835.
36. Steppel JH, Horton ES. Beta-cell failure in the pathogenesis of type 2 diabetes mellitus. *Curr Diab Rep* 2004; 4(3): 169-175.
37. Sermer M, Naylor CD, Gare DJ, Kenshole AB, Ritchie JW, Farine D et al. Impact of time since last meal on the gestational glucose challenge test: the Toronto Tri-Hospital Gestational Diabetes Project. *Am J Obstet Gynecol* 1994; 171(3): 607-616.
38. Metzger BE. The Pasadena Consensus Meeting: summary & highlights [Vortrag]. 5th International Symposium on Diabetes and Pregnancy; 26.-28.03.2009; Sorrent, Italien 26.03.2009.
39. Coustan DR, Lowe LP, Metzger BE. The hyperglycemia and adverse pregnancy outcome (HAPO) study: can we use the results as a basis for change? *J Matern Fetal Neonatal Med* 2010; 23(3): 204-209.
40. 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.
41. Lefebvre C, Manheimer E, Glanville J. Searching for studies. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. New York: Wiley; 2008. S. 95-150.

Anhang A: Suchstrategien

1. EMBASE

Suchoberfläche: Ovid

EMBASE 1980 to 2009 October 23

Es wurden folgender Filter übernommen: EMBASE-RCT-Strategy minimizing difference between sensitivity and specificity [40].

#	Searches
1	exp glucose tolerance test/
2	blood glucose monitoring/
3	glucose blood level/
4	insulin treatment/
5	insulin/dt
6	(glucose tolerance test or glucose test or ogtt).ti,ab.
7	((glucose or glycemi* or glycaemi*) adj3 (control* or treatment or manage* or monitor* or therapy)).ti,ab.
8	or/1-7 [OGTT]
9	pregnancy diabetes mellitus/
10	gestational.ti,ab.
11	or/9-10 [Gestationsdiabetes]
12	random*.tw.
13	clinical trial*.mp.
14	exp health care quality/
15	or/12-14 [Wong 2006, high sensitivity]
16	and/8,11,15

2. Medline

Suchoberfläche: Ovid

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1950 to Present

Es wurden folgende Filter übernommen: Cochrane HSSS sensitivity-maximizing version (2008 revision) [41].

#	Searches
1	Blood Glucose Self-Monitoring/
2	Glucose Tolerance Test/
3	Insulin/tu
4	(glucose tolerance test or glucose test or ogtt).ti,ab.
5	((glucose or glyce* or glycaemi*) adj3 (control* or treatment or manage* or monitor* or therapy)).ti,ab.
6	or/1-5 [OGTT]
7	Diabetes, Gestational/
8	Pregnancy in Diabetics/
9	gestational.ti,ab.
10	or/7-9 [Gestationsdiabetes]
11	randomized controlled trial.pt.
12	Randomized Controlled Trials as Topic/ [Ergänzung: Mesh 2009]
13	controlled clinical trial.pt.
14	Controlled Clinical Trials as Topic/ [Ergänzung: Mesh 2009]
15	randomized.ab.
16	placebo.ab.
17	drug therapy.fs.
18	randomly.ab.
19	trial.ab.
20	groups.ab.
21	or/11-20
22	(animals not (humans and animals)).sh.
23	21 not 22 [Cochrane HSSS sensitivity-maximizing version (2008 revision)]
24	and/6,10,23

Suchoberfläche: Pubmed

Search	Most Recent Queries
#8	Search #5 and (#6 or #7)
#7	Search publisher[sb]
#6	Search in process[sb]
#5	Search #3 AND #4
#4	Search gestational[tiab]
#3	Search (#2 OR #1)
#2	Search (glucose[tiab] or glyce*mi*[tiab] or glycaemi*[tiab]) AND (control*[tiab] or treatment[tiab] or manage*[tiab] or monitor*[tiab] or therapy[tiab])
#1	Search glucose tolerance test[tiab] or glucose test[tiab] or ogtt[tiab]

3. The Cochrane Library

Suchoberfläche: Wiley

Cochrane Central Register of Controlled Trials (Clinical Trials); Issue 4, 2009 of the Cochrane Library

ID	Search
#1	MeSH descriptor Blood Glucose Self-Monitoring explode all trees
#2	MeSH descriptor Glucose Tolerance Test explode all trees
#3	MeSH descriptor Insulin explode all trees with qualifier: TU
#4	(glucose tolerance test or glucose test or ogtt):ti,ab
#5	((glucose or glyce*mi* or glycaemi*) NEAR/3 (control* or treatment or manage* or monitor* or therapy)):ti,ab
#6	(#1 OR #2 OR #3 OR #4 OR #5)
#7	MeSH descriptor Diabetes, Gestational explode all trees
#8	MeSH descriptor Pregnancy in Diabetics explode all trees
#9	<u>gestational:ti,ab</u>
#10	(#7 OR #8 OR #9)
#11	(#6 AND #10)

Anhang B: Liste der im Volltext überprüften, aber ausgeschlossenen Studien mit Ausschlussgründen

Teilziel 1 „Screening“

keine

Teilziel 2 „Therapie“

Ausschlussgrund T3: keine Vergleichsgruppe wie unter 4.1.2 definiert

Moses RG, Barker M, Winter M, Petocz P, Brand-Miller JC. Can a low-glycemic index diet reduce the need for insulin in gestational diabetes mellitus? A randomized trial. *Diabetes Care*. 2009; 32(6): 996-1000.

Ausschlussgrund T4: kein Studientyp wie unter 4.1.5 definiert

Lundberg GD. Metformin trumps insulin in the treatment of gestational diabetes. *Medscape J Med* 2008; 10(7): 179.

Thomas J. Metformin safe treatment for gestational diabetes. *Australian Journal of Pharmacy* 2009; 90: 73.

Anhang C: Offenlegung potenzieller Interessenkonflikte des externen Reviewers

Im Folgenden sind die potenziellen Interessenkonflikte des externen Reviewers dargestellt. Alle Informationen beruhen auf Selbstangaben anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter <http://www.iqwig.de> abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externer Reviewer

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Horvath, Karl	nein	nein	ja	ja	ja	nein	ja	nein

Im „Formblatt zur Darlegung potenzieller Interessenkonflikte“ wurden folgende 8 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb der letzten 3 Jahre bei einer Person, Institution oder Firma¹ abhängig (angestellt) beschäftigt, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut² finanziell profitieren könnte?

Frage 2: Beraten Sie oder haben Sie innerhalb der letzten 3 Jahre eine Person, Institution oder Firma direkt oder indirekt³ beraten, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 3: Haben Sie abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre im Auftrag einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, Honorare für Vorträge, Stellungnahmen, Ausrichtung und oder Teilnahme an Kongressen und Seminaren - auch im

¹ Mit solchen „Personen, Institutionen, oder Firmen“ sind im Folgenden alle Einrichtungen gemeint, die direkt oder indirekt einen finanziellen oder geldwerten Vorteil aus dem Ergebnis Ihrer wissenschaftlichen Arbeit für das Institut ziehen könnten. Hierzu gehören z.B. auch medizinische Einrichtungen, die eine zu bewertende medizinische Intervention durchführen und hierdurch Einkünfte erhalten.

² Mit „wissenschaftlicher Arbeit für das Institut“ sind im Folgenden alle von Ihnen für das Institut erbrachten oder zu erbringenden Leistungen und/oder an das Institut gerichteten mündlichen und schriftlichen Recherchen, Bewertungen, Berichte und Stellungnahmen gemeint.

³ „Indirekt“ heißt in diesem Zusammenhang z.B. im Auftrag eines Institutes, das wiederum für eine entsprechende Person, Institution oder Firma tätig wird.

Rahmen von Fortbildungsveranstaltungen, oder für (populär-) wissenschaftliche oder sonstige Aussagen oder Artikel erhalten?⁴

Frage 4: Haben Sie und/oder die Einrichtung, für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre von einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und/oder die Einrichtung, bei der Sie angestellt bzw. beschäftigt sind innerhalb der letzten 3 Jahre sonstige finanzielle oder geldwerte Zuwendungen (z.B. Ausrüstung, Personal, Reisekostenunterstützung ohne wissenschaftliche Gegenleistung) von einer Person, Institution oder Firma erhalten, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile (auch in Fonds) von einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 7: Haben Sie persönliche Beziehungen zu einer Person, Firma oder Institution bzw. Mitarbeitern einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 8: Gibt es andere bislang nicht dargestellte potenzielle Interessenkonflikte, die in Beziehung zu Ihrer wissenschaftlichen Arbeit für das Institut stehen könnten?

⁴ Sofern Sie von einer Person, Institution oder Firma im Verlauf der letzten 3 Jahre mehrfach Honorare erhalten haben reicht es aus, diese für die jeweilige Art der Tätigkeit summiert anzugeben.