

# SWEPIS-Studie (SWEdish Post-term Induction Study): Geburtseinleitung in SSW 41 versus in SSW 42 bei risikoarmen Schwangerschaften

Wennerholm UB et al. Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (SWEdish Post-term Induction Study, SWEPIS): multicentre, open label, randomised, superiority trial. *BMJ* 2019; 367: l6131

## Zusammenfassung

Die hier zusammengefasste randomisierte kontrollierte Studie wurde vorzeitig beendet. Dies wirft viele Fragen auf. Daher gehen wir in unserem kritischen Kommentar stark auf die Planungsphase dieser Studie ein und besprechen das Outcome, das scheinbar den Abbruch notwendig werden ließ.

## Ziele

Ziel der 16 Forscherinnen und Forscher, mehrheitlich aus dem Karolinska Institut in Stockholm, Schweden, war es, die bislang in Schweden übliche Routine bei Terminüberschreitung (Einleitung 14-15 Tage nach errechnetem Termin) mit einer etwas früheren Einleitung (7-8 Tage nach errechnetem Termin) zu vergleichen und dabei sowohl kindliches als auch mütterliches Outcome zu berücksichtigen [1].

## Methodik

Vor Beginn der SWEPIS-Studie (SWEdish Post-term Induction Study) veröffentlichte das Forschungsteam im März 2015 das zuvor genehmigte Studienprotokoll [2]. Durch die Vorabveröffentlichung sollte verhindert werden, dass Änderungen während des Forschungszeitraums nach Belieben der Zuständigen erfolgen konnten – und damit das Ergebnis bewusst beeinflussbar wäre. Die Forschenden legten hierbei als primäres Outcome eine Kombination unter anderem folgender Befunde (composite perinatal outcome) fest.

- Totgeburt
- Tod des Kindes postnatal bis zum Alter von 27 Lebenstagen
- 5-Minuten Apgar-Wert < 7
- pH-Wert in der Nabelschnurarterie < 7
- hypoxisch-ischämische Enzephalopathie Grad 1-3
- intrakranielle Blutung
- Krampfanfälle
- Mekonium-Aspirations-Syndrom
- mechanische Beatmung innerhalb von 72 Stunden nach der Geburt
- Plexus-Brachialis-Verletzung im direkten Zusammenhang mit der Geburt.

Finanzielle Unterstützung kam von der Foundation of the Health and Medical care committee of the Region of Vastra Gotaland, Sweden, und vom Health Technology Centre der Schwedischen Regierung. Das Ethikvotum des Regional Ethics Board in Göteborg fiel positiv aus [2]. Zur Überwachung der Studie wurde ein unabhängiges Gremium (Data and Safety Monitoring Board) aus einem Statistiker (Hans Wedel) und zwei geburtshilflich Tätigen mit langer Berufserfahrung (Lars-Åke Mattson als Geburtshelfer und Elisabeth Jangsten als Hebamme) gebildet. Die Forschungsleitung berichtete alle Todesfälle und andere schwerwiegenden Ereignisse sofort nach Auftreten bzw. Meldung aus den Kliniken an diese drei Personen.

Im Vorfeld berechnete das Forschungsteam, wie viele Teilnehmerinnen benötigt werden, damit eine Reduktion des primären Outcomes um ein Drittel darstellbar ist. Anhand der Daten aller Schwangeren in der Schwangerschaftswoche 41 + 3 in der Region Skåne in den Jahren 2000 bis 2010 schätzten die Forscherinnen und Forscher das Auftreten des primären Outcomes zu diesem Zeitpunkt auf 2,74% aller Kinder. Obwohl Totgeburt und Tod des Kindes postnatal bis zum Alter von 27 Lebenstagen zwei verschiedene Komponenten des Outcomes sind, gibt das Forscherteam nur den gemeinsamen Wert von 0,53% für die Region Skåne an [2]. Zur Totgeburtenrate fehlt daher der Vergleichswert. Als Signifikanzniveau für den geplanten Vergleich legten die Forschenden den üblichen 0,05 Wert sowie eine Power von 80% fest und nahmen eine Dropout-Rate von 10% aller Teilnehmerinnen an. Mit diesen Vorgaben errechneten sie eine Stichprobengröße von 10.038 Teilnehmerinnen gleichmäßig verteilt auf beide Gruppen und legten vorab eine Zwischenauswertung aller bis dahin gewonnenen Daten für den Zeitpunkt fest, an dem Daten von 50% der angestrebten Geburtenanzahl vorliegen.

Ab September 2015 baten die Forschenden Low-Risk-Schwangere in 14 Kliniken mit angeschlossenen Zentren für Schwangerenvorsorge um Teilnahme an ihrer Studie. Die Einschlusskriterien der Teilnehmerinnen waren nicht einheitlich [1], auch wenn die Forschenden diesen Umstand in ihrer Publikation nicht als kritisch hervorheben: Der Geburtstermin wurde bei der Ultraschalluntersuchung entweder im ersten oder zweiten Trimester festgelegt. Die abwartende Betreuung verlief nach der jeweiligen Routine der 5 universitären und 9 ländlichen Zentren für Schwangerenbetreuung. Eine Kontrolluntersuchung erfolgte bei allen Schwangeren eine Woche nach dem ET, im Zeitraum von drei Tagen zwischen 40 SSW + 6 bis 41 SSW + 1. Dieser Untersuchungstermin war auch gleichzeitig der Zeitpunkt zur Randomisierung. Die allermeisten Frauen aus der Region Stockholm haben während dieser Untersuchung auch eine Ultraschalluntersuchung bei einer Hebamme oder einer ärztlichen Fachperson erhalten. Frauen mit einem auffälligen Befund schlossen die Forschenden von der Studienteilnahme aus. Obwohl die Ultraschalluntersuchung im Studienprotokoll vorgesehen war, war dieses Angebot in den 9 ländlichen Zentren nicht immer gegeben. Hier bekamen die interessierten Teilnehmerinnen von der zuständigen Forschungshebamme die jeweils übliche Routineuntersuchung mit Kontrolle des Blutdrucks, des Symphysen-Fundus-Abstandes und der kindlichen Herz-töne (Dopton). Ein Nachlassen der kindlichen Bewegungen, Verdacht auf Wachstumsretardierung oder Präeklampsie erforderte Handlungsbedarf und bedeutete einen Ausschluss dieser Frauen von der Teilnahme an der Studie. Bei allen anderen Frauen wurden die Geburten nach dem Zufallsprinzip entweder innerhalb 24 Stunden und damit 7-8 Tage nach errechnetem Termin (Studiengruppe) oder aber erst etwa eine Woche später, 14-15 Tage nach errechnetem Termin (Kontrollgruppe) eingeleitet.

## Ergebnisse

Von den insgesamt 12.554 Frauen, die die Voraussetzungen für eine Teilnahme erfüllten, lehnten 9.792 die Studienteilnahme ab. Nur 2.760 Frauen mit einer unkomplizierten Einlingschwangerschaft (22%), und damit weniger als die über 30%, die laut Studienprotokoll erwartet wurden, nahmen in den drei Jahren bis zum Abbruch Ende September 2018 an der Studie teil und wurden für eines der beiden Verfahren randomisiert. Die Gruppe mit der früheren Einleitung (7-8 Tage nach errechnetem Termin) hatte eine Gruppengröße von 1.381 Teilnehmerinnen und diejenige mit der Routinebehandlung eine Größe mit ähnlich vielen Teilnehmerinnen, 1.379, erreicht. Die schleppende Rekrutierung wird in einer Antwort der Forschenden im offenen Reviewprozess des Artikels mit der Überzeugung der Schwangeren beziehungsweise der Behandelnden begründet, dass Abwarten die bessere Option sei [3]. Relativ wenige Teilnehmerinnen (3,5% der Studiengruppe und 2% der Kontrollgruppe) haben nicht die Behandlung erhalten, die eigentlich für ihre Gruppe vorgesehen war. Entweder waren keine Kapazitäten in der betreffenden Klinik für eine Einleitung zum frühen Termin vorhanden oder die Frauen wollten eine Einleitung früher als geplant. 76 Geburten verliefen demnach nicht wie geplant, werden aber nach dem Prinzip „intention-to-treat“ in ihrer ursprünglichen Gruppe mit ausgewertet.

Von den Ergebnissen seien hier speziell diejenigen präsentiert, die zum Abbruch der Studie am 2. Oktober 2018 führten: Im Gegensatz zur Studiengruppe (7-8 Tage nach ET), in der kein kindlicher Todesfall verzeichnet wurde, traten 6 Fälle in der Kontrollgruppe (14-15 Tage nach ET) auf. Auch wenn die kindliche Sterblichkeit als Einzeloutcome zu den sekundären Outcomes zählt (und damit die Gruppengröße nicht daraufhin berechnet ist, einen signifikanten Unterschied in der kindlichen Sterblichkeit aufzeigen zu können), konnte es das Data and Safety Monitoring Board ethisch nicht vertreten, die Studie weiter

fortzuführen („it was not considered ethical to continue the study“).

## Schlussfolgerung der Autorinnen / Autoren

Die Forscherinnen und Forscher sehen ihre Ergebnisse trotz des Studienabbruchs als verwertbar an und schlussfolgern: Auch wenn die Ergebnisse mit Vorsicht zu bewerten sind, sollte Schwangeren eine Einleitung mit 41 SSW+0 Tage oder früher angeboten werden.

## Kommentar

Die Entscheidung, die Studie aus den vom Forscherteam genannten Gründen vorzeitig abzubrechen, ist insofern irritierend, als bis zum Abbruch die getestete Intervention gegenüber dem üblichen Vorgehen einen Benefit aufzeigte und damit nicht zu erwarten war, dass sich die perinatale Mortalität durch den Abbruch verbessern ließe. Es ist zudem bekannt, dass vorzeitige Studienabbrüche wegen eines Benefits den Nutzen einer Intervention tendenziell größer erscheinen lassen, als dies mit einer Durchführung der Untersuchung bis zur geplanten Stichprobengröße zu beweisen wäre [4]. Somit wurde hier ein Ergebnis angenommen (und als Empfehlung auch veröffentlicht), das sich bei einer Fortführung der Studie vielleicht gar nicht bestätigt hätte. Diese Kritik wird in einem Leserbrief von Ole Olsen aus dem Department für Public Health an der Universität von Kopenhagen vorgebracht [5].

Abgesehen von diesem Einwand lohnt es sich, die zum Abbruch der Studie führenden Todesfälle näher zu betrachten: Es handelt sich um 5 Totgeburten und einen Tod in der frühen Neugeborenenperiode (innerhalb der ersten 7 Lebenstage). Somit errechnet sich eine kindliche Todesrate von 0,4% (6/1.379). Diese Rate übersteigt demnach nicht die Rate von 0,53%, die als Referenzwert in der Region Skåne ermittelt wurde. Dennoch scheint die Meldung das dreiköpfige Gremium so sehr alarmiert zu haben, dass es nicht die vorgesehene Zwischenauswertung (bei Erreichen von 50 Prozent der 10.038 erforderlichen Geburten) abwarten, sondern bereits nach Erhebung von nur 27,5% der erforderlichen Geburten die Entscheidung zum Abbruch herbeiführen wollte.

Dem Review-Verfahren zu diesem Artikel kann entnommen werden, dass die Fälle erst nach der Entscheidung statistisch ausgewertet wurden. Das eigentlich zu untersuchende primäre, zusammengesetzte Outcome zwischen der Studiengruppe (Einleitung 7-8 Tage nach errechnetem Termin) mit 33 Fällen auf 1.381 Kinder (2,4%) und der Gruppe mit Routinebehandlung mit 31 Fällen auf 1.379 Kinder (2,2%) unterscheidet sich statistisch gesehen nicht, wie es das Relative Risiko von 1,06 mit einem 95%-Konfidenzintervall von 0,65–1,73 zeigt. Beide Prozentangaben liegen niedriger als die in der Region Skåne ermittelten Werte von 2,74% aller Kinder. Allerdings ist diese Berechnung nicht zuverlässig, da sie erst bei der Auswertung der 10.038 erforderlichen Geburten zuverlässige Aussagekraft gewonnen hätte.

Die Ursachen für die Todesfälle scheinen vielfältig: Ein Kind verstarb wohl intrauterin kurz nach der Aufnahme in die Klinik. Das Obduktionsergebnis zeigte einen (nicht lebensbedrohlichen) Herzfehler. Auch bei den anderen vier Totgeburten ließen sich keine sicheren Todesursachen nachweisen. Eines der Kinder war für sein Gestationsalter zu klein. Das postpartal verstorbene Kind hatte eine hypoxisch-ischämische Enzephalopathie. Alle 6 Todesfälle ereigneten sich bei Erstgebärenden, die nicht in der Stockholmer Region versorgt wurden und demnach eine andere, eventuell weniger ausführliche Schwangerenvorsorge erhielten als in dieser Region. Es bleibt unklar, in welcher Schwangerschaftswoche die betroffenen 6 Schwangeren bei Geburt waren, denn nur der Durchschnitt von 292 Tagen und das Minimum von 289 sowie das Maximum von 294 Tagen wird für die gesamte Kontrollgruppe angegeben. Es steht damit jedoch fest, dass keine Schwangere aus der gesamten Stichprobe den Zustand einer eigentlichen Übertragung erreicht hat. Auch der Zeitpunkt des Wehenbeginns (mit oder ohne Einleitung) und der Geburtsmodus werden für diese 6 Geburten nicht dargestellt. In der Gruppe mit Routinebehandlung hatte die Mehrheit, 66,7% der Frauen (920 von 1.379), einen spontanen Wehenbeginn, gegenüber 14,1% der Frauen mit der frühen Einleitung. Eine vaginale Geburt ohne Einsatz von Instrumenten erlebten in beiden Gruppen gleich viele Frauen: etwa 83%.

Der Peer-Review-Prozess des am 19. Juli 2019 beim British Medical Journal eingereichten Artikels ist offen und einsehbar [3]. Auf den ersten Entwurf hin, in dem die Forschenden den Abbruch der Studie noch mit dem statistisch signifikanten Unterschied in der perinatalen Mortalität zwischen den Gruppen begründen, schlägt der erste Reviewer als Erklärung stattdessen die Formulierung „Abbruch aus Sicherheitsgründen“ vor. Die zweite Reviewerin, Philippa Middleton, sieht keine Formfehler und äußert sich nur positiv über die Studie. An dieser Stelle sei bemerkt, dass Middleton die Hauptautorin des Cochrane Reviews „Induction of labour at or beyond 37 weeks' gestation“ ist.

Interessant gestaltet sich die Debatte, die auf der Leserbriefseite der Zeitschrift zu dieser Studie entbrennt [3]. Es steht der Vorwurf im Raum, dass die gesamte Studie jetzt nicht mehr nutzbar ist, da man sich nicht an die vorab festgelegte Vorgehensweise gehalten hat. Das mit dieser Kritik konfrontierte Forschungsteam entgegnete mit dem kaum mehr anfechtbaren Argument, welches der erste Reviewer als Erklärung für den Abbruch vorgeschlagen hatte: Die Studie wurde aus Sicherheitsgründen abgebrochen. Für diese Einschätzung gebe es keine statistische Grundlage und dieser Umstand hätte auch nicht im Studienprotokoll vorab festgelegt werden müssen. So der aktuelle Stand der Debatte.

Abgesehen von der Tatsache, dass das Outcome zum Abbruchzeitpunkt noch gar nicht schlechter war als die Anhaltswerte aus der Vergleichsregion Skåne und dass der Studienabbruch im Artikel zuerst nicht explizit mit der Sicherheit der Ungeborenen begründet wurde, ist es mehr als fraglich, ob die Studienzwischenergebnisse von dieser renommierten Zeitschrift in einer Form, als wäre es eine normale, wenn auch verkürzte Studie, hätten veröffentlicht werden dürfen.

Unser Hauptargument gegen die Schlussfolgerungen aus dieser abgebrochenen Studie ist jedoch das Folgende: Wir wissen nicht, wie viele andere Ungeborene gestorben sind,

bevor ihre Mütter randomisiert wurden. Bei wie vielen Schwangeren ist beim Randomisierungstermin ein intrauteriner Fruchttod festgestellt worden und damit eine Aufnahme in die Studie verhindert worden? Alle für die Studiengruppe randomisierten Frauen mit einer sofortigen Einleitung (oft am selben Tag, spätestens am Tag darauf) hatten nur wenige Stunden Zeit, von einem intrauterinen Todesfall getroffen zu werden – daher ist es nicht erstaunlich, dass keine Totgeburt verzeichnet wurde. Da es in der anderen Gruppe etwa eine Woche später zur Einleitung oder zum spontanen Geburtsbeginn kam und damit eine größere Zeitspanne für eine mögliche Totgeburt vorlag, wäre es nur logisch zu erheben, wie viele Totgeburten sich eine Woche vor der Randomisierung **bei allen Frauen** ereignet haben. Wenn wir aus diesem Grund die Rate der Totgeburten nicht vergleichen können und nur die nach der Geburt verstorbenen Kinder anschauen, hätte das Gremium eine Differenz von nur einem Todesfall sicher nicht als Alarmzeichen gewertet.

Betroffene (Schwangere oder Mütter) wurden in den Planungsprozess dieser Studie nicht involviert. Möglicherweise wäre ihnen dieser ungleiche Informationsgehalt der Daten aufgefallen. Umso bedauerlicher ist es, dass dieses Studienfragment als vollwertiger Datensatz in die aktualisierte Fassung des Cochrane Reviews „Induction of labour at or beyond 37 weeks' gestation“ eingeflossen ist, Hauptautorin: Philippa Middleton.

### Autorinnen / Autoren

**Dr. rer. medic. Christine Loytved** ist Hebamme und Gesundheitswissenschaftlerin. Sie arbeitet als Dozentin am Institut für Hebammen, Departement Gesundheit der Zürcher Hochschule für Angewandte Wissenschaften in Winterthur (Schweiz).

**Dorothea Zeeb MSc** ist Hebamme und Senior Lecturer im Bachelorstudiengang Hebammen an der Fachhochschule Salzburg (Österreich).

### Korrespondenzadresse

E-Mail: loyt@zhaw.ch

### Literatur

- [1] Wennerholm UB et al. Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (SWE-dish Post-term Induction Study, SWEPIs): multicentre, open label, randomised, superiority trial. *BMJ* 2019; 367: l6131.
- [2] Elden H, Hagberg H, Wessberg A, et al. Study protocol of SWEPIs a Swedish multicentre register based randomised controlled trial to compare induction of labour at 41 completed gestational weeks versus expectant management and induction at 42 completed gestational weeks. *BMC Pregnancy Childbirth* 2016;16:49. doi:10.1186/s12884-016-0836-9
- [3] Author response (18. September 2019). Access document: [https://www.bmj.com/sites/default/files/attachments/bmj-article/peer-history/author\\_response\\_18.9.19.pdf](https://www.bmj.com/sites/default/files/attachments/bmj-article/peer-history/author_response_18.9.19.pdf). Internet: <https://www.bmj.com/content/367/bmj.l6131/peer-reviewbw>. <https://www.bmj.com/content/367/bmj.l6131/peer-review>
- [4] Bassler, D., Montori V. M., Briel, M., Glasziou, P. & Guyatt, G. (2008) Early stopping of randomized clinical trials for overt efficacy is problematic. *Journal of clinical epidemiology* 61 (3), 241–246. doi:10.1016/j.jclinepi.2007.07.016
- [5] Olsen O. Leserbrief vom 30.1. 2020: Statistical analysis in SWEPIs trial is flawed. Internet: <https://www.bmj.com/content/367/bmj.l6131/rapid-responses>

### Bibliografie

Die Hebamme 2020; 33: 12–14

DOI 10.1055/a-1247-8648

ISSN 0932-8122

© 2020. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany'