DOCTORAL
THESIS
SAHLGRENSKA
ACADEMY

Management of late term pregnancy

Mårten Alkmark

SAHLGRENSKA ACADEMY
INSTITUTE OF CLINICAL SCIENCES

UNIVERSITY OF
GOTHENBURG

2021

Management of late term pregnancy | Mårten Alkmark

ISBN 978-91-8009-192-3 (PRINT)
ISBN 978-91-8009-193-0 (PDF)
http://hdl.handle.net/2077/67129

Printed by Stema Specialtryck AB, Borås

Management of late term pregnancy

Mårten Alkmark

Centre of Perinatal Medicine & Health
Department of Obstetrics and Gynaecology
Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg,
Gothenburg, Sweden

UNIVERSITY OF GOTHENBURG

Gothenburg 2021

Cover illustration: Jan Funke

Management of late term pregnancy

marten.alkmark@vgregion.se

ISBN 978-91-8009-192-3 (PRINT)
ISBN 978-91-8009-193-0 (PDF)
http://hdl.handle.net/2077/67129

Printed in Borås, Sweden 2021
Printed by Stema Specialtryck AB

To my family who supported me all the way

Management of late term pregnancy

Mårten Alkmark

Centre of Perinatal Medicine & Health
Department of Obstetrics and Gynaecology, Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Background: The optimal time point to intervene and induce labour in women with a low-risk pregnancy, in order to decrease perinatal adverse outcome, is up for debate. Some advocate for induction of labour (IOL) at 41 gestational weeks (GW) and others for expectant management (EM) until 42 GW.

Aim: To clarify, in women with a low-risk singleton pregnancy, if a policy of IOL at 41 GW (late term) compared with EM until 42 GW (postterm) was superior, in terms of neonatal and maternal outcomes, as well as health economic aspects. Furthermore, different methods of IOL in late term/postterm pregnancies were assessed.

Material and methods: *Paper I* was a Swedish multicentre register-based randomised controlled trial in women with a low-risk late term/postterm singleton pregnancy (n=2 760) comparing IOL at 41 GW with EM until 42 GW. The trial was conducted between May 2016 and October 2018. Primary outcome was a composite of perinatal mortality and morbidity. *Paper II* was a one-step individual participant data (IPD) meta-analysis (n=5 161 for aggregate data and 4 561 for IPD) and included trials comparing IOL in women with a low-risk singleton pregnancy at 41 GW with EM until 42 GW. Primary outcome was a composite of perinatal morbidity and mortality. Subgroup analysis was performed on maternal age, body mass index and parity. *Paper III* was a cost-effectiveness analysis alongside Paper I (n=2 746). Primary outcomes were costs per gained life year (LY) and quality adjusted life year (QALY). *Paper IV* was a prospective cohort study on efficacy, safety and women's childbirth experience of IOL with oral misoprostol (OM) (n=744) compared with transvaginal balloon catheter (TVBC) (n=469). Women included in Paper I, who needed cervical ripening for IOL, were assessed. Primary efficacy outcome was vaginal delivery within 24 hours, primary safety outcomes were a composite of neonatal mortality and morbidity and a composite of maternal mortality and morbidity. Women's childbirth experience was measured with the Childbirth Experience Questionnaire (CEQ 2.0).

Results: *Paper I:* The primary outcome did not differ between the groups. However, the trial was truncated early due to safety reasons because a significant decreased perinatal mortality was seen in the IOL group (p=0.03). Similar results in both groups were reported regarding mode of delivery and maternal adverse outcomes, except endometritis. *Paper II:* Three trials were eligible and two contributed with IPD. The primary outcome was significantly lower in the IOL group, relative risk (RR) 0.43 (95 % confidence interval [CI] 0.21 to 0.91), as were perinatal mortality, Peto odds ratio 0.21 (95 % CI 0.06 to 0.78) and admission to neonatal care ≥4 days, RR 0.52 (95 % CI 0.32 to 0.85). Similar results in both groups were reported regarding mode of delivery and maternal adverse outcomes. The primary outcome was significantly lower in the IOL group in nulliparous, but not in parous women. *Paper III*: The incremental cost-effectiveness ratio for IOL compared with EM was €545 per LY (95 % CI ranging from lower costs and better health outcomes [dominant] to €4 002) and €623 per QALY (95 % CI dominant to €4 586). *Paper IV:* Vaginal delivery within 24 hours was significantly lower in the OM group compared with the TVBC group, adjusted RR 0.76 (95 % CI 0.64; 0.89). Primary neonatal and maternal safety outcomes did not differ between groups. Women's childbirth experience was overall positive and similar in the groups.

Conclusion: There are medical benefits of IOL at 41 GW compared with EM until 42 GW and it is cost-effective. TVBC was slightly more effective regarding efficacy. Women approaching 41+0 GW should receive unbiased information regarding benefits and risks with IOL at 41+0 GW compared with EM in order to make an informed decision for herself and her unborn infant.

Keywords: Induction of labour, late term pregnancy, postterm pregnancy, stillbirth, perinatal mortality, cost-effectiveness

ISBN 978-91-8009-192-3 (PRINT)
ISBN 978-91-8009-193-0 (PDF)
http://hdl.handle.net/2077/67129

SAMMANFATTNING PÅ SVENSKA

I höginkomstländer är risken generellt mycket låg (i Sverige <0,5%) för att ett barn dör före, under förlossningen, eller som nyfödd, så kallad perinatal död. Även risken att barnet får skador i samband med förlossningen är låg. Det är dock känt att det sker en viss riskökning, både vad det gäller perinatal död och sjuklighet ju längre en graviditet pågår efter vecka 40. Riskerna för barnet ökar markant vid 42 veckor och noll dagar (≥42+0, så kallad överburenhet) och framåt. Cirka 20% av alla förlossningar startar spontant först vid 41+0 veckor eller senare och 6% när de blivit överburna. Igångsättning av förlossning görs för att försöka minska riskerna för barnet, men skulle kunna vara förenat med ökade risker för mamman. I Sverige idag sätts i regel förlossningen igång vid 42 graviditetsveckor hos friska kvinnor med en normal graviditet. I andra länder såsom Danmark, USA och Storbritannien erbjuds dessa kvinnor igångsättning från och med 41+0 veckor. I WHOs riktlinje från 2018 rekommenderar man igångsättning vid 41+0 veckor.

I våra studier ingick friska kvinnor med en normal enkelbördig graviditet, som nått 41 veckor av sin graviditet där förlossningen inte startat spontant. Syftet med studierna var att undersöka om igångsättning av förlossningen vid 41+0-2 veckor minskade perinatal död och sjuklighet, utan att öka hälsoriskerna för mamman eller påverka hälsoekonomin i negativ riktning jämfört med igångsättning vid 42+0-1 veckor. Vi undersökte även de två dominerande igångsättningsmetoderna som används i Sverige idag, med avseende på effektivitet, säkerhet och kvinnornas upplevelse.

I *delarbete I* gjordes ett slumpmässigt urval av 2 760 kvinnor till antingen igångsättning vid 41 veckor eller så avvaktades spontan förlossningsstart till 42 veckor då igångsättning utfördes. Det huvudsakliga utfallsmåttet var en kombination av perinatal död och sjuklighet bland barnen. Studien fick avbrytas i förtid, då det var signifikant färre dödsfall i gruppen som sattes igång vid 41 veckor. Man fann dock ingen skillnad mellan grupperna gällande det huvudsakliga kombinerade utfallsmåttet, död och sjuklighet. Man fann inte någon skillnad mellan grupperna gällande andel kejsarsnittsförlossningar, stora blödningar eller stora bristningar hos mamman. I *delarbete II* utfördes en systematisk översikt av litteraturen avseende vår huvudsakliga frågeställning i syfte att slå samman flera studier och därmed få ett större antal kvinnor som forskningsunderlag. Tre randomiserade studier med denna frågeställning fanns publicerade och två kunde bidra med forskningsdata på individnivå. Totalt ingick 4 561 kvinnor med data på individnivå. Vårt huvudsakliga utfallsmått var även här en kombination av perinatal död och sjuklighet hos barnen. Vi fann en statistiskt säkerställd lägre andel av kombinationen döda och sjuka

barn i gruppen som sattes igång vid 41 veckor jämfört med 42 veckor och detta gällde framfor allt förstföderskor. Vi såg även en statistiskt säkerställd skillnad avseende en lägre andel döda barn och barn i behov av neonatal vård fyra dagar eller längre. Dessa resultat uppnåddes utan att andelen kejsarsnitts-förlossningar, stora blödningar eller stora bristningar hos mamman ökade. Vi såg dessutom en signifikant minskning av andelen kvinnor med havandeskaps-förgiftning i gruppen som sattes igång vid 41 veckor. *Delarbete III* baserades på delarbete I och vi undersökte de hälsoekonomiska aspekterna av att ändra handläggningen till igångsättning vid 41 veckor jämfört med att vänta till 42 veckor. Den totala kostnaden (kostnad för öppenvård, förlossningsvård och eventuell neonatalvård) skilde sig inte signifikant mellan grupperna. I gruppen som sattes igång vid 41 fulla veckor var hälsovinsten större, än i den andra gruppen tack vare minskningen av dödsfall. Kostnaden för ett vunnet kvalitetsjusterat levnadsår (ett år i full hälsa) om man sätter igång förlossningen vid 41 veckor var 6 170 kr, vilket är långt under vad Socialstyrelsen anser vara en hög kostnad och därmed en begränsning för införandet av metoden. Deras gräns för hög kostnad ligger på ca 500 000 kr. *Delarbete IV* baserades också på delarbete I. Vi undersökte om det var några skillnader gällande effektivitet, säkerhet och kvinnors upplevelse mellan att sätta igång förlossningen, hos kvinnor med en omogen livmodertapp, med en tablettbehandling eller med en mekanisk metod bestående av en ballongkateter som placeras vid livmodertappens inre öppning. Vi fann att igångsättning med den mekaniska metoden resulterade i något fler vaginala förlossningar inom 24 timmar från igångsättningens start. Inga säkerställda skillnader gällande säkerhetsprofilen för varken barnet eller mamman kunde påvisas. Vi kunde inte heller finna några skillnader i kvinnors upplevelse mellan metoderna. Samtliga kvinnor rapporterade en generellt god förlossningsupplevelse.

Sammanfattningsvis så finns det medicinska fördelar med att sätta igång förlossningen vid 41+0-2 veckor, istället för att avvakta spontan förlossningsstart till 42+0-1 veckor. Den hälsoekonomiska utvärderingen visar också att igångsättning vid 41 veckor är kostnadseffektiv, dvs medför en låg kostnad per vunnet kvalitetsjusterat levnadsår. Igångsättning av förlossningen bör erbjudas alla kvinnor – särskilt förstföderskor – vid 41 veckor. Kvinnor som inte är förlösta när de närmar sig 41 veckor bör få tydlig, både skriftlig och muntlig, information kring fördelar och nackdelar med igångsättning av förlossningen så att de själva kan göra ett informerat val.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. Wennerholm UB, Saltvedt S, Wessberg A, Alkmark M, Bergh C, Brismar Wendel S, Fadl H, Jonsson M, Ladfors L, Sengpiel V, Wesström J, Wennergren G, Wikström AK, Elden H, Stephansson O, Hagberg H. Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (SWEdish Postterm Induction Study, SWEPIS): multicentre, open label, randomised superiority trial. *BMJ* (Clinical research ed). 2019;367:l6131.

II. Alkmark M, Keulen J.K.J, Kortekaas J.C, Bergh C, van Dillen J, Duijnhoven R.G, Hagberg H, Mol B.W, Molin M, van der Post J.A.M, Saltvedt S, Wikström AK, Wennerholm UB, de Miranda E. Induction of labour at 41 weeks or expectant management until 42 weeks: a systematic review and an individual participant data meta-analysis of randomised trials. *PLoS Medicine* 2020;17(12):e1003436.

III. Alkmark M, Wennerholm UB, Saltvedt S, Bergh C, Carlsson Y, Elden H, Fadl H, Jonsson M, Ladfors L, Sengpiel V, Wesström J, Hagberg H, Svensson M. Induction of labour at 41 weeks of gestation versus expectant management and induction of labour at 42 weeks of gestation: A cost-effectiveness analysis. *Submitted, pending decision.*

IV. Alkmark M, Carlsson Y, Brismar Wendel S, Elden E, Fadl H, Jonsson J, Ladfors L, Saltvedt S, Sengpiel V, Wessberg A, Wikström AK, Hagberg H, Wennerholm UB. Efficacy and safety of oral misoprostol versus transvaginal balloon catheter for labor induction: An observational study within the SWEdish Postterm Induction Study (SWEPIS). *Under revision.*

CONTENT

ABBREVIATIONS

ACOG	American College of Obstetricians and Gynaecologists
AD	aggregate data
Adj	adjusted
AFI	amniotic fluid index
aOR	adjusted odds ratio
ARRIVE	A Randomized Trial of Induction Versus Expectant Management
BMI	body mass index
BPD	biparietal diameter
CD	caesarean delivery
CEA	cost-effectiveness analysis
CEQ 2.0	Childbirth Experience Questionnaire
CI	confidence interval
CNOGF	National College of French Obstetrician and Gynaecologists
CPP	cost per patient
CRL	crown rump length
CTG	cardiotocography
DSOG	Danish Society of Obstetrics and Gynaecology
eCRF	electronic case report form
EDD	estimated date of delivery
EM	expectant management

EMR	electronic medical record
EQ-5D	The EuroQol -5 Dimension measure
Gbg	Gothenburg
GRADE	The Grading of Recommendations Assessment, Development and Evaluation
GW	gestational week
HIE	hypoxic ischaemic encephalopathy
HELLP	Haemolysis, elevated liver enzymes and low platelet count
ICD	international classification of disease
ICER	incremental cost effectiveness ratio
INDEX	INDuction of labour at 41 weeks versus a policy of EXpectant management until 42 weeks
IOL	induction of labour
IPD	individual participant data
IQR	interquartile range
ITT	intention to treat
LY	life year
MA	meta-analysis
MAS	meconium aspiration syndrome
MBR	Medical Birth Register
MD	mean difference
n	numbers

NA	not applicable
NE	not estimated
NGF	Norwegian Society of Gynaecology
NICE	National Institute for Health and Care Excellence
NICU	neonatal intensive care unit
NNT	number needed to treat
OM	oral misoprostol
OR	odds ratio
PNM	perinatal mortality
PS	propensity score
QALY	quality adjusted life years
RCT	randomised controlled trial
RD	risk difference
Ref	reference
RR	relative risk
SALAR	The Swedish Association of Local Authorities and Regions
SD	standard deviation
SNQ	Swedish Neonatal Quality Register
SOGC	The Society of Obstetricians and Gynaecologist of Canada
SPR	Swedish Pregnancy Register
Sthlm	Stockholm

SWEPIS	SWEdish Post-term Induction Study
TVBC	transvaginal balloon catheter
VAS	Visual Analogue Scale
vs	versus
WHO	The World Health Organization

DEFINITIONS IN SHORT

Certainty of evidence according to GRADE	The Grading of Recommendations Assessment, Development and Evaluation (short GRADE) is a tool for grading quality (or certainty) of evidence and strength of recommendations. It is considered the standard in evaluation certainty of evidence
High certainty	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty	Confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

INTRODUCTION

Definition of term, late term and postterm pregnancy

The length of a pregnancy in a population is usually based on large epidemiological studies. The World Health Organization (WHO) defines a normal pregnancy to be 280 days (40 weeks and zero days, 40+0 gestational weeks [GW]) (Timonen et al., 1965). The estimation of pregnancy length of 280 days (nine months) is much older than the 1960's and Hermann Boerhaave (1668–1738) was the first to propose a rule on how to calculate the estimated date of delivery (EDD) and Franz Carl Naegele (1778-1851) clarified the rule; adding nine months and seven days to the first day of the last menstrual period (Lawson, 2020). In Sweden, up to this date, a pregnancy is considered to be 280 days. However, in modern times, due to the entrance of ultrasound assessment of EDD, the length of pregnancy has been altered to 283 days in e.g. Norway (Kessler et al., 2019).

In 1977, WHO defined a term pregnancy lasting between 259 and 293 days, i.e. between 37 weeks and zero days and 41 weeks and six days (37+0 to 41+6 GW), from the first day of the last normal menstrual period (WHO, 1977). The definition was based on the assumption of a homogeneous and favourable perinatal outcome during this period of pregnancy (ACOG, 2013). In 2013, American College of Obstetricians and Gynaecologists (ACOG) published a consensus document of an updated definition of a term pregnancy (ACOG, 2013). The new definition divided term pregnancy into early term; a pregnancy between 37+0 and 38+6 GW, full term; a pregnancy between 39+0 and 40+6 GW and late term; a pregnancy between 41+0 and 41+6 GW (ACOG, 2013). The reason for the division into these three categories was based on research that argued for a more heterogeneous perinatal outcome during the term period than previously stated and full term was the period with the most favourable perinatal outcome (Reddy et al., 2011; Tita et al., 2011).

A pregnancy lasting 294 days and more (≥42+0 GW) from the first day of the last normal menstrual period is defined as a postterm pregnancy according to WHO (WHO, 1977) and the update by ACOG did not change this definition (ACOG, 2013). Prolonged pregnancy is also used in the literature instead of postterm pregnancy. However, it is sometimes used for pregnancies lasting 42+0 GW or more and sometimes for pregnancies lasting 41+0 GW or more.

In this thesis the nomenclature late term (41+0-41+6 GW) and postterm pregnancy (≥ 42+0 GW) will be used.

Dating the pregnancy

In order to know when a pregnancy is full term, late term or postterm properly dating is needed. Historically determining the EDD was made by defining the first day of the last normal menstrual period. In 1812 Franz Carl Naegele published a suggestion how to calculate EDD by adding a year, subtracting three months and adding 7 days to the last normal menstrual period, which later on was called Naegele's rule. This rule is based on menstrual cycles of 28 days and assumes the pregnancy to be 280 days.

Ultrasound assessment was introduced in obstetrics during 1970's and during 1990's it was widely used to estimate EDD because it showed to be more accurate than dating according to the last menstrual period (Geirsson et al., 1991; Tunón et al., 1996). The ultrasound examination was carried out between 18 and 20 GW in the beginning, but in 2017, ACOG published a committee opinion statement recommending estimating the EDD with ultrasound during the first trimester up to 13+6 GW (ACOG, 2017). The procedure of estimating EDD with ultrasound includes two different measurements; crown rump length (CRL) (H. P. Robinson, 1973; H. P. Robinson et al., 1975) or biparietal diameter (BPD) measurement (Campbell, 1969; Donald et al., 1961; Selbing et al., 1985).

In Sweden today, most pregnant women are offered an ultrasound between 12+4 and 13+6 GW for a first assessment and estimation of EDD and a second ultrasound between 18 and 20 GW to further assess the structural anatomy of the fetus (SFOG-råd, 2019). In Sweden in 2019, estimation of EDD with BPD was performed between 12+4 and 13+6 GW in 49.9 % of pregnancies, in the second trimester in 42.4 % and with CRL during an early ultrasound before 10 GW in 7.7 % (Register, 2020). The BPD should be ≥21 mm in order to be the base for dating the pregnancy (SFOG-råd, 2019). Estimation of EDD is carried out during the second ultrasound if BPD <21mm. According to the Swedish Pregnancy Register (SPR) report from 2019 estimating EDD with measurement of BPD with ultrasound between 12+4 and 13+6 GW results in

an estimation of EDD 0.4 days earlier than the actual birth with a 95 % confidence interval (CI) of 0.3 to 0.5 and a standard deviation (SD) of 7.7 days compared to dating between 18 and 20 GW with BPD (EDD 0 days earlier than the actual birth; CI -0.1 to 0.1, SD 8.0 days) or dating with CRL (EDD 1.1 days earlier; 95 % CI 0.9 to 1.3, SD 7.7) (Register, 2020). To be noted is that, as stated previously, there are several studies (Bergsjo et al., 1989; Kessler et al., 2019; Tunón et al., 1996) indicating that the mean pregnancy duration probably is longer than the internationally acknowledged 280 days, which might be the explanation for that dating between 12+4 and 13+6 GW results in a difference of 0.4 days.

In pregnancies where assisted reproductive technology is used, for example in vitro fertilization, estimation of EDD is based on the day of embryo transfer and the age of the embryo as recommended in Swedish guidelines (SFOG-råd, 2019).

Incidence of late term and postterm pregnancies

The incidence of postterm pregnancies in a population is dependent on the dating method in combination with the characteristics of the population. In 2013, the incidence of pregnancies 41+0 GW and beyond varied between 3.5 % (Malta) and 25.7 % (Denmark). The corresponding numbers for postterm pregnancies was 0.1 % (Malta) and 6.7 % (Sweden) (Seijmonsbergen-Schermers et al., 2020).

As stated above, the recommended time to estimate EDD is between 12+4 and 13+6 GW (SFOG-råd, 2019). A Cochrane review from 2015 (Whitworth et al., 2015) concluded that estimation of EDD by ultrasound in the first trimester may result in fewer inductions of labour (IOL) due to postterm pregnancies compared to an estimation of EDD with ultrasound in the second trimester or if the latest normal menstrual period was used. In Sweden, in 2019, 4.8 % of a population turned postterm when EDD was estimated between 12+4 and 13+6 GW, compared with 5.2 % in a population where the estimation was performed during 18-20 GW and 5.8 % in the group estimated with CRL (Register, 2020).

Other factors that affect the incidence of postterm pregnancies are fetal surveillance with ultrasound in term pregnancies, ratio between nulliparous and parous women and rate of obesity and maternal age (Norwitz et al., 2007; Roos et al., 2010; Shea et al., 1998). In populations where fetal surveillance with ultrasound between 40+0 and 41+6 GW is performed an increased number of interventions such as IOL and caesarean deliveries due to suspicion of placental insufficiency will be carried out, hence, decreasing the rate of postterm pregnancies (Norwitz et al., 2007). In a Swedish observational study from 2010 a more than 50 % increased risk for postterm pregnancy was found for nulliparous compared with parous women and women ≥35 years old compared with women between 20 and 24 years of age (Roos et al., 2010). Furthermore, the risk of postterm delivery increased gradually with an increase in body mass index (BMI) (BMI 20-24 as reference) and the risk was increased by 60 % if you had a BMI ≥30 (Olesen et al., 2006; Roos et al., 2010). Moreover, the rate of preterm birth and rate of scheduled caesarean deliveries also affects the postterm delivery rate.

Before estimation of EDD was performed with ultrasound the incidence of postterm deliveries in high-income countries varied between four and 11 % (Shea et al., 1998). In the United States the incidence of pregnancies 41+0 GW and beyond and postterm pregnancies were 18 % and 10 % respectively in 1998 and in 2005 14 % and six percent, respectively (Norwitz et al., 2007). The incidence of postterm pregnancy in Sweden was 8.4 % between 1982 and 1991 (Ingemarsson et al., 1997) as compared to 5.4 % in 2017 (MBR, 2020). This decrease was seen even though other characteristics in the population such as increasing BMI and maternal age would suggest an increase in postterm pregnancies. In addition, about 18-20 % of the Swedish pregnant population reaches 41+0 GW and beyond. During 2018, the incidence of pregnancies 41+0 GW and beyond in the Västra Götaland region, comprising of approximately 20 % of deliveries in Sweden, was 22 % (data from the Obstetrix data base in Västra Götaland).

Aetiology of postterm pregnancy

The aetiology of postterm pregnancy is not well known and the most common reason for postterm pregnancy is inaccuracy in estimating the EDD. However, risk factors such as nulliparity, maternal age ≥30 years, and obesity are recognized, as mentioned earlier, and might affect the 'true' aetiology of postterm pregnancy (Arrowsmith et al., 2011; Kortekaas et al., 2015; Olesen et al., 2006; Roos et al., 2010). Genetic factors also play a role in postterm pregnancies. Thus, a woman who herself, was born postterm has a 49 % increased risk of a postterm pregnancy compared with a women born between 28 and 41 GW. The risk is 23 % higher if the father of the child was born postterm compared with fathers born between 28 and 41 GW (Morken et al., 2011). In addition, male fetal sex has also been associated with postterm pregnancy (Divon et al., 2002). Whether male fetuses are more prone to be postterm or if it is due to size differences between male and female fetuses when EDD is determined leading to misclassification of gestational age is unknown (Skalkidou et al., 2010).

The physiological process of onset of labour is not fully known. Some rare disorders associated with postterm pregnancy are anencephaly, trisomy 16 and 18 and Seckel's dwarfism (primordial dwarfism characterized by microcephaly, typical craniofacial appearance with a beak-like prominent nose, mental retardation, and various other congenital anomalies (Takikawa et al., 2008)) (Shea et al., 1998). Conditions affecting the adrenal-pituitary function is also associated with postterm delivery (Norwitz et al., 2007). A deficiency in placental surfactant is another rare condition associated with postterm pregnancy (ACOG, 2014) as is extra uterine pregnancy that continues to full term pregnancy (Shea et al., 1998).

Fetal/neonatal risks in late and postterm pregnancies

Stillbirth and neonatal mortality

The risk of stillbirth increases with gestational age after 32 GW and approximately half of all fetal deaths after 32 GW occur at term (Reddy et al., 2006). During 2016, 433 stillbirths (≥22+0 GW) occurred out of 121 511

newborns (3.6/1 000) in Sweden and approximately 25 % were born ≥39 GW (SOS, 2018). In two large national register studies, one in Sweden (1982-1991) and one in Denmark (1978-1993), the risk of stillbirth was significantly higher in postterm pregnancies than in term pregnancies (Ingemarsson et al., 1997; Olesen et al., 2003). In the Swedish study they found an increased rate of stillbirth in nulliparous women between 41+0 and 41+6 GW (1.86/1 000) and between 42+0 and 42+6 GW (2.26/1 000) as compared with gestational age between 40+0 and 40+6 GW (1.23/1 000) (Ingemarsson et al., 1997). The Danish study showed an adjusted odds ratio (aOR) of 1.36 (95 % CI 1.08–1.72) for stillbirth in postterm compared to term pregnancies (Olesen et al., 2003). In both the Swedish and the Danish study, the calculations were based on the number of women delivered in each GW. If you instead calculate with the number of fetuses at risk (deliveries and ongoing pregnancies in each GW) as denominator, the incidence of stillbirth will change. In the Swedish study, in 41 and 42 GW, as compared to 40 GW the stillbirth risk increased even more for each GW. In nulliparous women the incidence was 0.62/1 000, 1.26/1 000 and 2.26/1 000, in 40, 41, and 42 GW respectively (Cnattingius et al., 1998). Both studies were carried out before estimation of EDD with ultrasound as standard. Hence, both studies are limited by the different estimation methods of the EDD (ultrasound, last normal menstrual period, examination of the size of the uterus by doctor in early pregnancy, or the 'best' estimate based on the methods mentioned above).

A recent cohort study from Denmark, conducted between the years 2007 and 2018, calculated the cumulative risk of stillbirth between 41+0 and 41+6 GW. An exponential increase in stillbirth from 0.16/1 000 at 41+0 GW to 1.25/1 000 at 41+6 GW was reported (Figure 1) (Lidegaard Ø 2020).

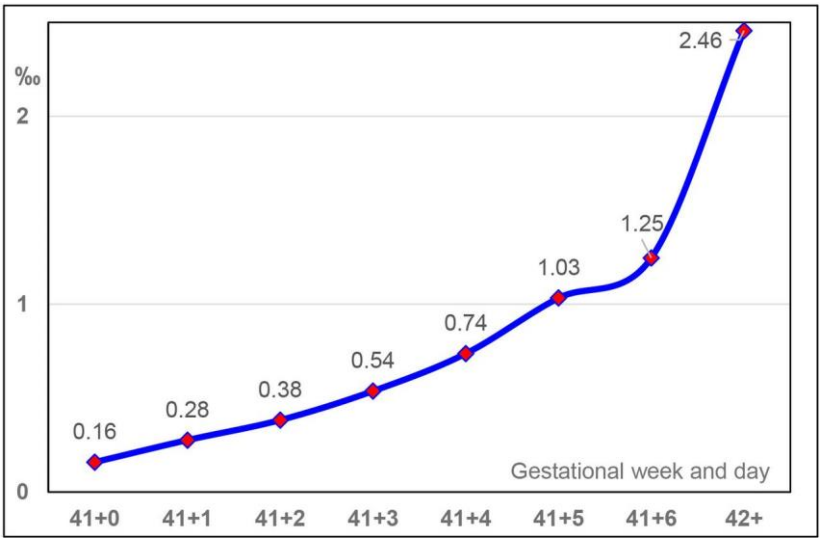

Figure 1. *Cumulated risk of intrauterine death from 41+0 gestational weeks per 1 000 ongoing pregnancies. Denmark 2007–2018. (Lidegaard Ø 2020)*

In a large cohort study from California they stratified the risk of stillbirth by gestational age in non-anomalous, term and postterm deliveries between 1997 and 2006 (n=3,820,826) (Rosenstein et al., 2012). EDD was based on latest normal menstrual period. The incidence of stillbirth for each GW (37 to 42 GW) was calculated as the number of stillbirths at that GW divided by the number of deliveries and ongoing pregnancies at that week i.e. gestation-week-specific risk. The incidence of stillbirth in 41 GW was 0.6/1 000 and in 42 GW 1.1/1 000 (Figure 2). The risk of stillbirth increased 2.0 times in 40 GW, 2.9 times in 41 GW and 5.1 times in 42 GW compared with 37 GW (Rosenstein et al., 2012).

The Swedish National Board of Health and Welfare published a report on stillbirth based on the birth cohorts from 2008 to 2016 in Sweden (SOS, 2018). An incidence of stillbirth between 41+0 and 41+6 GW of 2.0/1 000 was reported and the same incidence between 42+0 and 42+6 GW was seen. The overall incidence of stillbirth in Sweden during 2019 was 3.2/1 000 births (SOS, 2020). In May 2019, a meta-analysis (MA) of cohort studies published

between 1990 and 2017 comprising 15 million pregnancies in high-income countries was published (Muglu et al., 2019).

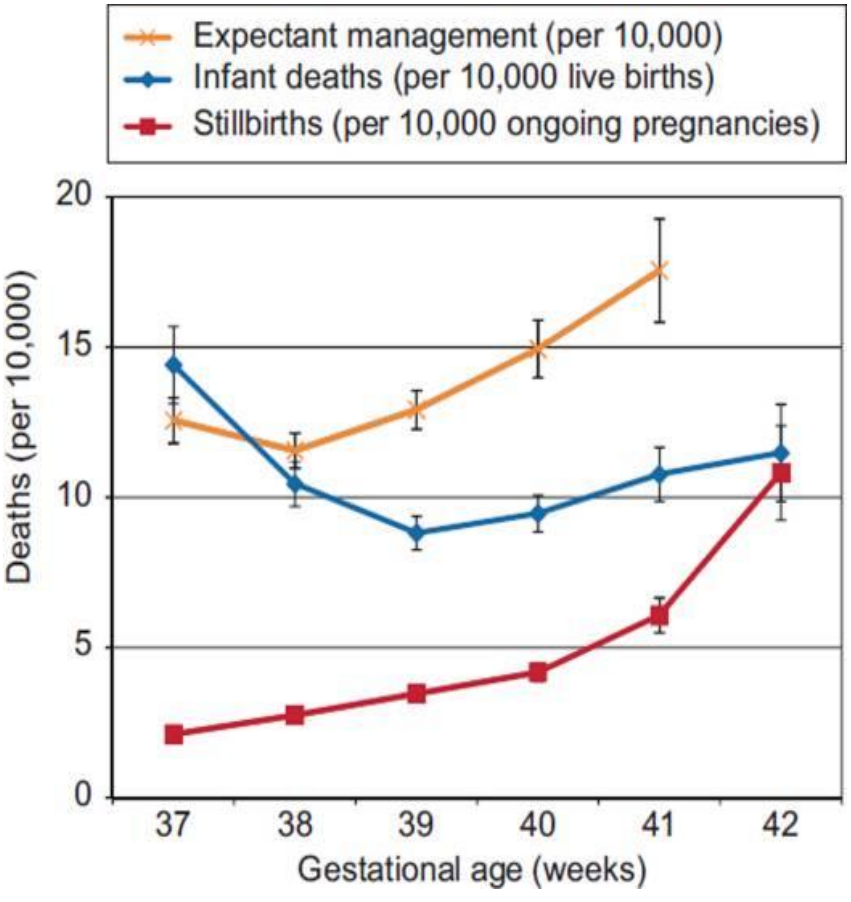

Figure 2. *This graph compares the risk of delivery (represented by infant death, i.e. death within the first year of life) with the risk of expectant management for 1 week (represented by the stillbirth rate plus the infant death risk at the subsequent gestational age) at each gestational age at term. The stillbirth rate also is displayed graphically to demonstrate its exponential rate of change. (Rosenstein et al., 2012)*

The analysis validated the increased risk of stillbirth after 37+0 GW and the relative risk (RR) increased significantly from 0.11/1 000 pregnancies (95 % CI 0.07-0.15) to 3.18/1 000 at 42 GW (95 % CI 1.84-4.35). The neonatal mortality remained the same between 38+0 and 41+0 GW, but increased after 41+0 GW (RR 1.87, 95 % CI 1.07 to 2.86) (Figure 3) (Muglu et al., 2019).

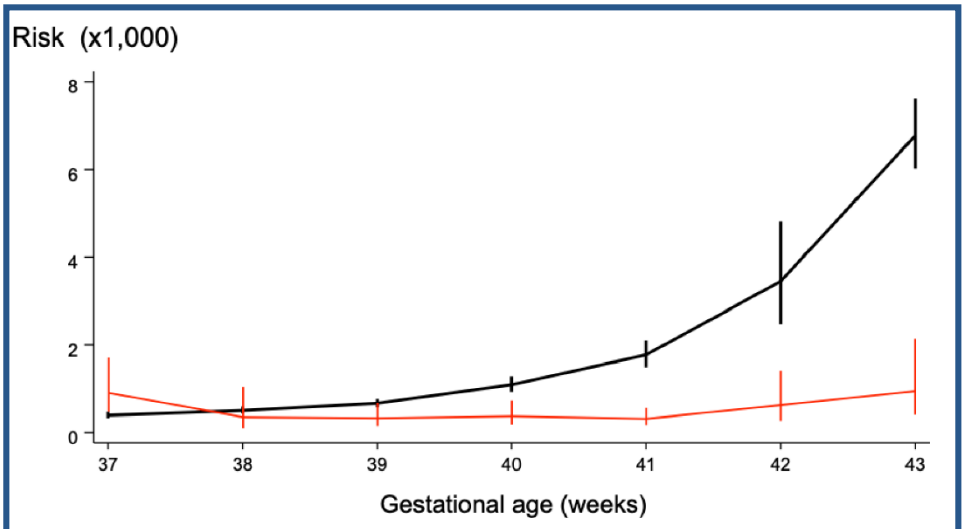

Figure 3. *Prospective risk of stillbirth per 1,000 pregnancies and risk of neonatal death per 1,000 deliveries by gestational age in pregnancies continued to term. Stillbirth risk (solid back line); neonatal death risk (solid red line) (Muglu et al., 2019)*

One article in a series on stillbirth published in Lancet reported 14 % of stillbirths in the world to be attributable to prolonged pregnancy (prolonged pregnancy not defined in the article) (Lawn et al., 2016). However, Flenady et al., 2011, reported population attributable risk of postterm pregnancy (the incidence of stillbirth attributed to postterm pregnancy) for stillbirth to be 0.3 % (Flenady et al., 2011). This rate was based on studies with a very low rate of postterm pregnancy of 0.9 %. Hence, the population attributable risk would be considerably increased in populations where IOL is not offered before 42+0 GW. Furthermore, in a report by the Swedish National Board of Health and Welfare in 2018, they estimated that approximately 1 % of all

stillbirths between 2008 and 2016 in Sweden occurred after 40+6 GW (SOS, 2018).

In summary, the proportion of late term and postterm pregnancy contributing to the incidence of stillbirth is uncertain.

Aetiology of stillbirth in late and postterm pregnancies

In low- and middle-income countries, overall contributors to stillbirth includes maternal malnutrition, maternal infectious disease, poorly managed hypertensive disorders of pregnancy and poor access to midwife led antenatal care and the option of caesarean delivery (Lawn et al., 2011; Lawn et al., 2016). In high-income countries a malfunctioning placenta resulting in intrauterine growth restriction or placental abruption is reported to be a major contributor to stillbirth (Flenady et al., 2011). However, the cause of stillbirth is not known in up to 70 % of stillbirths and it applies especially to stillbirth in term and postterm pregnancy (Flenady et al., 2011).

Distinguishing between a cause and a risk factor for stillbirth is not always possible due to knowledge gaps. Moreover, there are difficulties when attempting to assess and compare causes of stillbirth across different countries due to a variety of different classification systems (Aminu et al., 2017). As mentioned above, malfunctioning placenta is a contributor to stillbirth and according to a systematic review and MA on risk factors for stillbirth (Flenady et al., 2011), about 38 % of all stillbirths/neonatal deaths are related to a malfunctioning placenta (23 % were related to small for gestational age fetuses and 15 % were associated to placental abruption). However, there is evidence that <20 % of stillbirths in postterm pregnancies are associated with growth restriction defined as small for gestational age (small for gestational age defined as ≤2 standard deviations, according to Swedish sex specific reference (Marsal et al., 1996)) (Divon et al., 1998).

Fetal malformation is reported to account for less than 10 % of all stillbirths in most countries and obesity, hypertension and diabetes are estimated to contribute to approximately 10 % of the stillbirth rate in the world (Lawn et al., 2016). The corresponding figures for late term and postterm pregnancies are not known. In addition, both Flenady et al. and Lawn et al. recognise low

socioeconomic status, advanced maternal age and smoking as important preventable risk factors for stillbirth (Flenady et al., 2011; Lawn et al., 2016)

Non preventable risk factors for stillbirth are reported to be a nulliparous women and carrying a male fetus (Flenady et al., 2011). However, when Cnattingius et al. recalculated the incidence of stillbirth in the Swedish cohort study by Ingemarsson et al. with the approach of fetuses at risk parous women had an incidence of 0.73/1 000 at 40 GW, 1.00/1 000 at 41 GW and 1.51/1 000 at 42 GW indicating an increased risk with advancing gestational age (Cnattingius et al., 1998)

In summary, a decreasing placenta function is a possible aetiology of stillbirth in late term and postterm pregnancies. Furthermore, several risk factors are known, yet, for most cases of stillbirth in late term and postterm pregnancies the aetiology is unknown.

Neonatal adverse outcomes

Observational studies show that the risk of adverse perinatal outcomes, such as asphyxia, umbilical cord complications, meconium aspiration syndrome (MAS), sepsis, shoulder dystocia, traumatic injuries, pneumonia, neonatal convulsions and peripheral nerve damage in low-risk pregnancies gradually increases after 40 GW and is significantly elevated in postterm pregnancies (Linder et al., 2017; Olesen et al., 2003). There is also evidence that the risk of neonatal encephalopathy increases significantly in newborns born late term and postterm (41 GW: aOR 3.3 and GW 42: aOR 13.2 versus 39 GW) (Badawi et al., 1998). Furthermore, a Swedish cohort study reported a significantly increased risk of MAS and Apgar score <7 at five minutes in postterm nulliparous women, but not parous women, indicating that newborns of nulliparous women are at risk in postterm pregnancy (Lindegren et al., 2017).

In addition, a higher rate of developmental delay, behavioural, emotional and neuropsychiatric problems have been reported as long-term effects of being born postterm (El Marroun et al., 2012; Lindstrom et al., 2005). A Swedish study (n=354 newborns born postterm and a control group of 379 newborns born term) implicated an increased risk of developmental delay at the age of 4-4.5 years if born postterm compared to being born at term (OR 2.20; 95 % CI: 1.29-3.85) (Lindstrom et al., 2005). A study from the Netherlands (cohort

of 5 145 newborns with 382 newborns born postterm) assessed behavioural and emotional problems using a validated checklist (Child Behaviour Checklist, CBCL/1.5-5) at 18 and 36 months and reported an overall higher risk of behavioural problems if born postterm compared to term (OR 2.10, 95 % CI 1.32 to 3.36) (El Marroun et al., 2012). Children born postterm were also more likely to have attention deficit hyperactivity disorder symptoms (OR 2.44, 95 % CI 1.38 to 4.32).

In summary, according to observational studies and systematic reviews and MA, there is an increased risk of stillbirth, neonatal mortality and infant mortality when a pregnancy turns late term and postterm compared with full term. Furthermore, also according to observational studies, newborns delivered late-term and postterm are at greater risk of neonatal morbidity as well as adverse neuropsychiatric, behavioural and emotional long term effects.

Maternal risks in late term and postterm pregnancies

Complications for the mother also increase in pregnancies at 40 GW and beyond. A Danish register study demonstrated that the risk of labour dystocia, cephalo-pelvic disproportion, cervical lacerations, emergency caesarean deliveries, postpartum bleeding and puerperal infections was higher in postterm than in term pregnancies (aOR 1.2–1.6) (Olesen et al., 2003). Linder et al., also reported postterm pregnancy to be an independent risk factor for an increase in caesarean delivery and operative vaginal delivery (Linder et al., 2017). A retrospective cohort study from the US of low-risk pregnancies reported a significant increased risk of caesarean delivery due to non-reassuring fetal heart rate from 13.7 % at 39 GW to 23.5 % at 41 GW and 27.5 % at 42 GW (Caughey et al., 2007). Further, they reported an increase in perineal lacerations III and IV, chorioamnionitis, prolonged labour, postpartum haemorrhage and endometritis.

Women's experience

There are few studies exploring women's experience of late term and postterm pregnancy. A Norwegian randomised controlled trial (RCT) comparing IOL at 41 GW with expectant management until 43 GW asked the participants if they would prefer the same management in the next pregnancy. The trial included 508 women and 98 % completed the childbirth experience part of the trial. In the IOL group 74 % compared with 38 % of the women in the expectant management group would prefer the same management again (Heimstad, Romundstad, et al., 2007). Furthermore. In a trial comparing IOL at 39+0-6 GW with expectant management until 41+0-42+0GW in women ≥35 years old comprising 619 women, 83 % completed a questionnaire regarding their childbirth experience. The childbirth experience was similar in both groups (Walker et al., 2016). In addition, in an evaluation of women's childbirth experience in relation to IOL at 41 GW compared with expectant management until 42 GW Nilvér at al. showed an overall favourable experience with no significant between-group differences (Nilvér et al., 2021, Women's childbirth experience in the Swedish Postterm Induction Study [SWEPIS]: a multicentre, randomized, controlled trial, revision submitted Dec 2020).

In a systematic review of qualitative studies (eight studies included) in women's experience of childbirth in relation to postterm IOL they identified three different findings. Firstly, being subject of IOL due to postterm pregnancy meant that the expectation of spontaneous onset of labour had to be abandoned involving a shift in expectations. Additionally, women did not feel a part of the decision for IOL. Finally, the IOL process was perceived to be more focused on women fitting in to the delivery unit's organisation than being flexible to the pregnant woman's needs (Lou et al., 2019). Furthermore, more recent qualitative research by Wessberg et al. explored women's experience of not being delivered before 41 GWs (Wessberg et al., 2017, 2020). Women described the time from EDD to delivery as being in limbo and their negative feelings increased with gestational age (Wessberg et al., 2017). In both studies by Wessberg et al. women described a decline in trust in the body's ability to give birth, when the pregnancy exceeded 41 GW (Wessberg et al., 2017, 2020).

In summary, risk of adverse maternal outcome increases when the pregnancy turns postterm compared to full term pregnancies. Furthermore, women that reached 41 GW describes it as being in an indeterminate state and postterm

women report a lack of trust in their ability to give birth. However, their childbirth experience does not seem to be affected by a IOL or expectant management policy.

Prevention of postterm pregnancy

One of the most important methods of reduceing the rate of postterm pregnancy is to establish the gestational age as accurately as possible. Hence, as mentioned earlier, an accurate estimation of EDD with first trimester ultrasound decreases the incidence of postterm deliveries in a population. Furthermore, preventive public health work in order to decrease obesity among pregnant women could also decrease the incidence of postterm pregnancies.

In addition, there are several home remedies in different cultures in order to start the delivery when full term. Examples of methods include unprotected coitus (in order to get a prostaglandin effect in the cervix from the prostaglandins present in semen) and physical exercise (Norwitz et al., 2007). Further, castor oil has been suggested to have a ripening effect on the cervix, as have acupuncture and acupressure, but evidence is low for a favourable effect of these methods (Kelly et al., 2013; Smith et al., 2017). Sweeping the membranes has been studied more extensively and a recent Cochrane review has been published (Finucane et al., 2020). They conclude, with low certainty of evidence according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) (G. H. Guyatt et al., 2008), that sweeping the membranes might increase the probability of spontaneous onset of labour and reduce the risk of formal IOL. However, timing of the procedure and how many times to sweep the membranes is still uncertain. Furthermore, sweeping the membranes is associated with increased risk of vaginal bleeding and discomfort. A contraindication for the procedure is placenta praevia, which sometimes is undetected during pregnancy. Sweeping the membranes in a woman with undetected placenta praevia may result in heavy bleeding in an outpatient setting (ACOG, 2014). The relationship between group B streptococcus colonisation and sweeping the membranes is not known. If it could result in more chorioamnionitis or neonatal infections is not clear. Thus, how to council women with group B streptococcus colonisation regarding membrane sweeping is uncertain.

Management of late term pregnancies

Fetal surveillance

One possible way to reduce the risk of stillbirth and adverse neonatal outcome in late term and postterm pregnancies is to detect fetuses at risk. Signs of a fetus at risk are e.g., intrauterine growth restricted fetuses or decreased volume of amniotic fluid, both being a sign of a malfunctioning placenta. Another sign of a fetus at risk is decreased fetal movements (Rayburn, 1987). Fetal surveillance methods used, in late term low-risk pregnancies, for detecting fetuses at risk, are maternal registration of fetal movement, non-stress test and/or contraction/oxytocin stress test with cardiotocography (CTG), an ultrasound assessment of the fetal biophysical profile (heart rate, measurement of amniotic fluid, fetal breathing movements and fetal tone), a combination of non-stress test and amniotic fluid measurement, estimation of the fetal weight and Doppler assessment of blood flow in the umbilical artery (Alfirevic et al., 2015; Lalor et al., 2008). A single estimation of the fetal weight will only detect fetuses small for gestational age and not fetuses that are growth restricted. In order to detect a growth restricted fetus at least two ultrasound assessments of fetal weight two weeks apart is needed. In a Cochrane overview of Cochrane reviews evaluating antenatal interventions for reducing stillbirth, perinatal mortality and fetal loss they concluded that there are some interventions with clear benefits. However, most of them are applicable to a low- and middle-income country setting (balanced energy/protein intake, midwifery led care and trained traditional birth attendants) (Ota et al., 2020). The recommended frequency of surveillance varies and the optimal combination of method and frequency is up to debate.

Evidence that antepartum fetal surveillance in late term and postterm pregnancies reduces perinatal morbidity and mortality is weak. The evidence that maternal registration of fetal movements reduces perinatal adverse outcomes is conflicting (Bhatia et al., 2019; Mangesi et al., 2015; J. E. Norman et al., 2018).

Fetal surveillance with CTG has shown to be associated with high false-negative and false-positive results and without a clear effect on perinatal outcome in both high and low risk pregnancies (Evertson et al., 1979; Grivell et al., 2015; Manning et al., 1980; Schifrin, 1979).

A Cochrane review from 2008 comparing biophysical profile with non-stress test and/or the combination of non-stress test and measurement of amniotic fluid in high-risk pregnancies, including postterm pregnancies, concluded that there is insufficient evidence for the use of biophysical profile assessment as a test of fetal wellbeing compared to CTG or the combination of CTG and amniotic fluid assessment (Lalor et al., 2008). This review did not include any trials assessing late term pregnancies but included one trial assessing low-risk postterm pregnancies (n=145) showing no significant differences in perinatal outcome between a modified biophysical profile (computerised CTG, amniotic fluid index, and assessment of fetal breathing, tone and gross body movements) and standard CTG and measurement of the single deepest pocket of amniotic fluid (Alfirevic et al., 1995). The other four trials included a mixture of high-risk pregnancies. Using Doppler ultrasound in order to detect fetuses at risk in low-risk pregnancies lacks conclusive evidence of reduction in perinatal adverse outcome (Alfirevic et al., 2015). In contrast, using Doppler ultrasound in high-risk pregnancies was associated with a decrease in perinatal mortality (Alfirevic et al., 2017).

Furthermore, studies assessing detection rate of intrauterine growth restriction at term with routine ultrasound during the third trimester report conflicting results (Henrichs et al., 2019; Sovio et al., 2015). A cohort study from the UK reported a threefold increase in detection rate of small for gestational age fetuses. The routine ultrasound assessment at 28 and 36 GW had a sensitivity of 57 % compared with 20 % when indicated ultrasound assessment was performed (Sovio et al., 2015). In the population identified as potentially small for gestational age with routine ultrasound assessments an increased relative risk of 1.6 (95 % CI 1.22 to 2.09) for neonatal morbidity was reported. However, the rate of false positive results indicates that for every correctly identified small for gestational age infant two false positive results were present. Hence, the benefit gained from identification of small for gestational age fetuses needs to be balanced against the potential harm of unnecessary interventions due to false positive results. Moreover, they did not compare the routine ultrasound with indicated ultrasound regarding perinatal outcome. In a more recent stepped wedged cluster randomised trial from the Netherlands (Henrichs et al., 2019) they showed an increase in detection rate of small for gestational age fetuses with routine ultrasound assessment at 28-30 and 34-36 GW as compared with ultrasound assessment on indication from 19 % to

32 %. However, no difference in perinatal outcome could be detected. In addition, a Swedish cohort study from the Stockholm region compared a routine ultrasound assessment at 41+0 GW with ultrasound assessment on indication (Lindqvist et al., 2014). The routine assessment included a biophysical profile and measurement of fetal abdominal diameter. One birth centre performed ultrasound assessment on indication and another performed routine ultrasound assessments. Baseline characteristics differed significantly between the two centres. They reported an approximately two-fold increase in adverse neonatal outcome in the whole small for gestational age group (both centres) compared with average for gestational age newborns. The detection rate of fetuses being small for gestational age was higher at the routine ultrasound centre, but they could, however, not show a significant difference in perinatal outcome including perinatal mortality between the centres (Lindqvist et al., 2014). Ota et al. concludes that for fetal surveillance, in both high and low-risk pregnancies, with CTG or ultrasound assessment there are unknown benefits or harm or no effect or equivalence. An exception was computerised assessment of antenatal CTG compared to non-computerised antenatal CTG where a reduction in perinatal mortality was found (Ota et al., 2020).

In summary, today, there is no evidence that current fetal surveillance methods during late term and postterm pregnancy reduce perinatal morbidity and mortality.

Induction of labour

Another approach to reduce the risks of late term and postterm pregnancy is IOL. However, historically IOL has been associated with increased morbidity for the mother e.g., from increased rate of caesarean delivery due to failed inductions and operative vaginal deliveries compared with a spontaneous start of labour (Luthy et al., 2004; Roos et al., 2010; Vahratian et al., 2005). Thus, the balance when to induce labour in order to do good and not harm has been an important issue and research has been interpreted differently in the world. Hence, the time point when to induce late term/postterm pregnancies differs from country to country, region to region and even hospital to hospital. In Sweden, the general practice has been to induce labour at 42+0 GW, as in some

other countries. Different international guidelines on IOL in late term and postterm pregnancies are summarised in Table 1.

Table 1. *Summary of WHO and several international guidelines on management of late term and postterm pregnancy*

Guideline Country Year of publication	Recommendation on induction of labour	Recommendation on fetal surveillance
WHO **2018**	IOL is recommended for women who are known with certainty to have reached 41+0 GW	NA
ACOG **USA** **2014**	IOL after 42+0 GW and by 42+6 GW is recommended, given evidence of an increase in perinatal morbidity and mortality. IOL between 41+0 GW and 42+0 GW can be considered	Initiation of antepartum fetal surveillance ≥41+0 GW weeks of gestation may be indicated
CNOGF **France** **2013**	In the absence of a specific disorder, IOL can be proposed between 41+0 and 42 +6 GW	CTG twice or three times a week > 41+0 GW. Ultrasound assessment of AFI is recommended
DSOG **Denmark** **2011**	IOL should be initiated at a time point so women have given birth before 42+0 GW Women with increased risk for stillbirth include women with advanced age and obese women. These women are offered IOL at 41+0 GW depending on the number and severity of risk factors Women with prior caesarean delivery should be either induced or delivered by caesarean delivery at 41+3 GW	If the woman wants to wait for a spontaneous birth, antenatal monitoring twice weekly >41+3 GW with CTG, ultrasound and AFI assessment. If possible membrane sweeping is recommended
NGF **Norway** **2020**	In low-risk pregnancies, IOL is recommended between 42+0 and 42+2 GW All women should be offered routine ultrasound assessment at 41+0 GW. IOL is recommended at 41+0 GW if: 1. Estimated fetal weight <5th percentile 2. Reduced amniotic fluid volume (AFI <5 cm and/or deepest pocket <2 cm) 3. EDD estimated with ultrasound >14 days later than latest normal menstrual period 4. Mother's age> 38 years	After 41+0 GW ultrasound assessing AFI and CTG at 2-3-day intervals. IOL on indication. The woman's wish on IOL must be taken into account, but IOL must primarily be done after assessment by a specialist and on a medical basis
NICE **UK** **2008**	Women with uncomplicated pregnancies should usually be offered IOL between 41+0 and 42+0 GW to avoid the risks of prolonged pregnancy The exact timing should take into account the woman's preferences and local circumstances	From 42 GW, women who decline IOL should be offered increased antenatal monitoring consisting of at least twice-weekly CTG and ultrasound estimation of maximum amniotic pool depth
SOGC **Canada** **2017**	Women should be offered IOL at 41+0 to 42+0 GW, as the present evidence reveals a decrease in PNM without increased risk of Caesarean section	Antenatal testing used in the monitoring of the 41- to 42-GW pregnancy should include at least a CTG and an assessment of AFI

ACOG, American College of Obstetricians and Gynecologists; AFI, amniotic fluid index; CNOGF National College of French Obstetricians and Gynecologists, CTG, cardiotocography; DSOG, Danish Society of Obstetrics and Gynaecology; EDD, estimated date of delivery; GW, gestational week; IOL, induction of labour; NA, not applicable: NGF, Norwegian Society of Gynaecology and Obstetrics; NICE, National Institute for Health and Care Excellence; SOGC, The Society of Obstetricians and Gynaecologist of Canada; WHO, World Health Organization

However, more recent research has challenged the view that IOL increases maternal morbidity, more specifically the risk of caesarean and operative vaginal deliveries. A large retrospective cohort study, showed an increased risk for caesarean delivery if induced at 38+0, 39+0 or 40+0 GW compared with spontaneous onset of labour or induction later on. Yet, at 41+0 GW the risk of caesarean delivery was similar in the induced women and the women with spontaneous onset of labour or induction later on (Caughey et al., 2006). Several meta-analyses (MA) of RCTs in term pregnancies have found that IOL decreases the risk of caesarean delivery (Middleton et al., 2020; Mishanina et al., 2014; Wood et al., 2014). In one of them, the recent Cochrane review comparing IOL at 37+0 GW or beyond with expectant management without upper limit in gestational age, showed a slight reduction in caesarean delivery in the IOL group compared to the expectant management group (RR 0.90, 95 % CI 0.85 to 0.95; 31 trials, 21,030 women; moderate certainty of evidence) (Middleton et al., 2020). One of the largest trials included in the Cochrane review was Hannah et al., 1995. According to the study protocol IOL was performed with prostaglandin E2 in the IOL group but in the expectant management group IOL was only performed with amniotomy and oxytocin administration. If this was not possible a caesarean delivery was performed. This might be an explanation for the difference in caesarean delivery rates in the trial. Furthermore, it might have affected the total result of the aggregate analysis in the Cochrane review, hence, it was this trial and the ARRIVE trial (Grobman et al., 2018), that contributed with most data. The ARRIVE trial, compared IOL in nulliparous uncomplicated pregnancies at 39+0-4 GW with expectant management until 41+0 GW and reported a decrease in caesarean delivery in the IOL group compared with the expectant management group (18.6 % versus. 22.2 %, RR, 0.84; 95 % CI, 0.76 to 0.93) (Grobman et al., 2018). In the INDuction of labour at 41 weeks versus a policy of EXpectant management until 42 weeks (INDEX) trial, a non-inferiority RCT consisting of 1 801 low-risk pregnancies comparing IOL at 41+0-2 GW with expectant management and induction at 42+0 GW, there was no difference in the rate of caesarean delivery between the IOL and expectant management group (Keulen et al., 2019). In terms of operative vaginal delivery, the evidence is inconsistent. In the Cochrane review the risk of operative vaginal delivery was not different in the IOL versus expectant management group (Middleton et al., 2020). However, a decrease in operative vaginal deliveries with a IOL policy has been reported in some large observational studies (Stock et al., 2012; Zizzo

et al., 2017), and an increase in some (Cheng et al., 2012; Lindegren et al., 2017).

Several RCTs have been conducted on IOL at term and beyond exploring perinatal outcomes including perinatal morbidity, but most of them have too small sample sizes and diverse comparison groups. There were only two RCTs comparing induction at 41+0-2 GW with expectant management until 42+0 GW in the Cochrane review from 2018 (Gelisen et al., 2005; Keulen et al., 2019).

The Turkish study (Gelisen et al., 2005) (n=600) enrolled low-risk singleton pregnancies in cephalic presentation. Primary outcome was caesarean delivery rate, length of hospital stay and neonatal morbidity. A decrease in meconium-stained amniotic fluid and MAS in the IOL group was found. The caesarean delivery rate was similar in both groups. There was one stillbirth in the expectant management group.

The INDEX trial (Keulen et al., 2019), a non-inferiority trial (n=1801) included women with a singleton pregnancy in cephalic presentation. The primary outcome was a composite of perinatal mortality and neonatal morbidity. The research group could not show non-inferiority for expectant management compared with IOL. Hence, the composite perinatal outcome was significantly reduced in the IOL group. There were two stillbirths in the expectant management group and one in the induction group. There were no differences in maternal outcomes (Keulen et al., 2019).

In order to retrieve more information several systematic reviews and MA have been performed (Table 2).

The most recent Cochrane review (includes SWEPIS), which included trials with IOL as early as 37+0 GW and expectant management without no upper limit (includes 34 RCTs and over 21 000 women), concluded that IOL at or beyond term is associated with fewer perinatal deaths (RR 0.31, 95 % CI 0.15 to 0.64; 22 trials, 18 795 newborns, high certainty evidence). They also present lower rates of caesarean deliveries in the IOL group, as stated previously, compared with expectant management with a moderate certainty of evidence of evidence according to the GRADE system (G. Guyatt et al., 2011). Hence, the conclusion from the Cochrane review was that further research is needed

to establish the optimal time point for IOL in pregnancies at and beyond term (Middleton et al., 2020). In contrast to most systematic reviews addressing this subject Rydahl et al., in a systematic review evaluating IOL at 41 GW compared to expectant management, (including two randomised trials and two quasi-experimental trials, with a total of 5 119 women, and three cohort studies with 356 338 women) report an increase in caesarean delivery rate (RR 1.11, 95 % CI 1.09 to 1.14). According to the authors the review lacked statistical power to draw any conclusion regarding perinatal mortality.

Table 2. *Summary of systematic reviews and meta-analyses: Induction of labour at ≥41 weeks versus expectant management. Outcome: Perinatal mortality, meconium aspiration syndrome, admission to neonatal intensive care unit and caesarean delivery*

Study author, year of publication	No of RCTs	RR (95 % CI)	RR (95 % CI)	RR (95 % CI)	RR (95 % CI)
Myers 2002	15	Advantage induction	No difference	NA	No difference
Sanchez-Ramos 2003	16	3/3 159 vs 10/3 067 0.41 (0.14-1.18)	6/752 vs 12/666 0.46 (0.18-1.21)	291/2 495 vs 313/2 510 0.92 (0.78-1.10)	661/3 292 vs 709/3 216 0.88 (0.78-0.99)
Gulmezoglu 2006	19	1/2 986 vs 9/2 953 0.30 (0.09-0.99)	12/860 vs 31/853 0.39 (0.21-0.75)	41 GW 548/2 512 vs 294/2493 RR not calculated 42 GW 23/210 vs 24/212 RR not calculated	41 GW 559/2 883 vs 630/2 872 0.92 (0.76-1.12) 42 GW 110/407 vs 111/403 0.97 (0.72-1.31)
Wennerholm 2009	13	1/3 119 vs 3/3 097 0.33 (0.10-1.09)	14/1 114 vs 33/1 107 0.43 (0.23-0.79)	279/2 766 vs 312/2 747 0.89(0.77-1.03)	658/3 318 vs 750/3 299 0.87 (0.80-0.96)
Hussain 2011	16	1/3 315 vs 3/3 282 0.31 (0.11-0.88)	14 /1 114 vs 33/1 107 0.43 (0.23-0.79)	NA	NA
Gulmezoglu 2012	22	>41 GW 1/2 814 vs 9/2 785 0.30 (0.09-0.99)	≥41 GW 33/1 189 vs 66/1 182 0.50 (0.34-0.73)	>41 GW 289/2 676 vs 321/2 659 0.90 (0.78-1.04)	>41 GW 591/3 004 vs 649/2 990 0.91 (0.82-1.00)
Wennerholm 2012	19	41 GW 0/3 164 vs 8/3 137 0.28 (0.08-1.00)	41 GW 8/917 vs 24/916 0.39 (0.18-0.88)	41 GW 279/2 766 vs 312/2 747 0.89 (0.77-1.04)	41 GW 594/3 212 vs 670/3 201 0.91 (0.78-1.07)
Middleton 2018	30	≥41 GW 2/4 217 vs 13/4 191 0.33 (0.13-0.87)	133/3 887 vs 173/3 894 0.77 (0.62-0.97)	≥41 GW 307/3 704 vs 350/3 693 0.88 (0.76-1.01	≥41 GW 774/4 407 vs 857/4 407 0.90 (0.83-0.98)
Rydahl, Eriksen 2019	4	1/2 582 vs 6/2 527 0.22 (0.04-1.32)	11/554 vs 17/554 0.65 (0.31-1.37)	64/2 371 vs 66/2 300 0.95 (0.68-1.33)	420/2 582 vs 376/2 527 1.11 (0.98-1.26)

Table 2. Summary of systematic reviews and meta-analyses: Induction of labour at ≥41 weeks versus expectant management. Outcome: Perinatal mortality, meconium aspiration syndrome, admission to neonatal intensive care unit and caesarean delivery. ***Continued***

Study author, year of publication	No of RCTs	RR (95 % CI)	RR (95 % CI)	RR (95 % CI)	RR (95 % CI)
Alkmark, Berglin 2020	3	1/2 581 vs 9/2 580 0.20 (0.06-0.70)	6/2 581 vs 17/2 580 0.38 (0.17-0.86)	130/2 581 vs 165/2 580 0.79 (0.63-0.99)	298/2 581 vs 311/2 580 0.96 (0.82-1.11)
Middleton 2020	34	≥41 GW 2/5 472 vs 19/5 437 0.26 (0.11-0.64)	152/8 325 vs 202/8 297 Peto odds ratio 0.75 (0.62-0.92)	≥41 GW 359/4 958 vs 425/4 932 0.84 (0.74-0.96)	≥41 GW 882/5 662 vs 979/5 642 0.90 (0.83-0.97)

CD, caesarean delivery; CI, confidence interval; GW, gestational weeks; MAS, meconium aspiration syndrome; NA, not available: NICU, neonatal intensive care unit; PNM, perinatal mortality; RCT, randomised controlled trial; RR, risk ratio; vs, versus

In addition, several cohort studies have been performed. Table 3 summarises six Nordic large cohort studies with conflicting results from a systematic literature search (Alkmark, Berglin, et al., 2020).

Table 3. *Nordic observational studies with data from Medical Birth Registries analysing the effect of induction of labour in late term and postterm pregnancies on perinatal mortality and caesarean delivery. Results from a systematic literature search*

Author Year of publication Country	Study Period	Study population	Outcome		Comment
			Perinatal mortality	Caesarean delivery	
Grunewald 2011 Sweden	2000-2007	n=119 198 (total) ≥41+3 GW Three study groups ≥41+3 according to rate of deliveries ≥42+3 GW 2000-2004: n=27 311, n=13 160, n=33 206 2005-2007: n=16 865, n=7822, n=20 834 Sthlm formed a separate group	No difference in PNM, a reduction in Sthlm 2005-2007 (5.9 % ≥42+3 GW) vs 2000-2004 (21.0 % ≥42+3 GW) PNM: AOR 0.52: 95 % CI 0.31-0.83)	Rates of CD did not change in Sthlm during the two time periods	No national guidelines Routines differed: IOL at 42 +0 or 43+0 GW. Sthlm changed routines in 2005 from IOL at 43+0 GW to IOL at 42+0 GW
Lindegren 2017 Sweden	2001-2013	n=199 770 (total) ≥41+3 GW. Three groups according to rate of deliveries ≥42+3 GW and parity Nulliparous: n=35 133, n=33 177, n=35 465 Parous: n=31 230, n=31 621, n=33146	EM vs most active management: AOR not significant different for nulliparous or parous women	EM vs most active management: AOR (95 % CI) Nulliparous: 0.82 (0.78-0.86) Parous: 0.85 (0.79-0.91)	No national guidelines during study period. Routines differed: IOL at 42 +0 or 43+0 GW. Those with the most active management had, in nulliparous women, less women with MAS and Apgar>7 at five minutes
Pyykonen 2018 Finland	2006-2012	n=212 716 (total) IOL: Group 1 40+0-40+2 GW n=6882 Group 2 40+3-40+5 GW n=5543 Group 3 40+6-41+1 GW n=5115 Group 4 41+2-41+4 GW n=5581 Group 5 41+5-42+0 GW n=10 167	RR (95 % CI) Group 3: 1.00 (0.06-15.98) Group 4: 2.00 (0.18-22.05) Group 5: 2.50 (0.78-7.97)	RR (95 % CI) Group 3: CD 1.17 (1.06-1.28) Group 4: 1.19 (1.09-1.29) Group 5: 1.01 (0.94-1.07)	Policy of IOL at 42+0 to 42+2. PS matched control groups of equal size. Each group compared to all births beyond the studied GA period and the spontaneous births during the studied GA period

Table 3**. Nordic observational studies with data from Medical Birth Registries analysing the effect of induction of labour in late term and postterm pregnancies on perinatal mortality and caesarean delivery. Results from a systematic literature search.* ***Continued

Author Year of publication Country	Study Period	Study population	Outcome		Comment
			Perinatal mortality	Caesarean delivery	
Zizzo 2017 Denmark	2012-2014 and 2008-2010	n=87 505 (total) 2012-2014 IOL 41+2-41+6 GW (n=42 075) vs 2008-2010 IOL ≥42+0 GW (n=45 430)	2012-2014 vs 2008-2010: AOR (95 % CI) PNM: 0.62 (0.39-0.96) Stillbirths: 0.50 (0.29-0.89)	AOR (95 % CI) 0.98 (0.94-1.02)	National guidelines in Denmark changed in 2011 from IOL at ≥42+0 to IOL at 41+2-41+6 GW
Rydahl, Declercq 2019 Denmark	2000-2010 and 2012-2016	n=152 887 (total) ≥41+3 GW and onwards	No significant difference between a predicted PNM without change in policy and the actual PNM after change in policy	No significant difference before and after policy change in the rate of caesarean deliveries	National guidelines in Denmark changed in 2011 from IOL at ≥42+0 to IOL at 41+2-41+6 GW ITSA was performed Predicted stillbirth rate with and without policy change was compared
Lidegaard 2020) Denmark	2007-2018	n=179 734 (total) ≥41+0 GW	PNM shows an inverse correlation with increase in IOL at the time for policy change. A decrease from 1.3/1 000 in 2007/2008 to 0.38/1 000 in 2011/2012	NA	National guidelines in Denmark changed in 2011 from IOL at ≥42+0 to IOL at 41+2-41+6 GW

AOR, adjusted odds ratio; AS, Apgar score; CI, confidence interval; EM, expectant management; GA, gestational age; GW, weeks of gestation; IOL, induction of labour; ITSA, Interrupted Time Series Analysis; MAS, meconium aspiration syndrome; NA, not applicable; PNM, perinatal mortality; PS, propensity score; RR, relative risk; Sthlm, Stockholm; vs, versus**;**

The publications from Zizzo et al., Rydahl et al. and Lidegaard et al. all examined the same shift in policy regarding IOL in late and postterm pregnancies in Denmark, but within different timeframes (Lidegaard Ø 2020; Rydahl, Declercq, et al., 2019; Zizzo et al., 2017). Lidegaard et al. presents an inverse correlation between the rate of IOL in the population and the rate of stillbirth (Figure 4 and 5).

In contrast, Rydahl et al. presented an interrupted time series analysis of almost the same population and concluded that there was no difference in perinatal or stillbirth before and after the policy change (Figure 6). However, the predicted time series analysis seems to have missed the increase in stillbirth between 2011-2012 and 2015-2016. In Rydahl et al. they reported the stillbirth rate each year but if it was below five the actual number was not revealed due to confidentiality reasons. Despite the fact that the tendency for the increase is visible in the table in the publication, the authors does not comment on it.

In summary, before we started our trial presented in Paper I the evidence for IOL at 41+0 GW improving perinatal outcome compared with expectant management until 42+0-1 GW was still unclear.

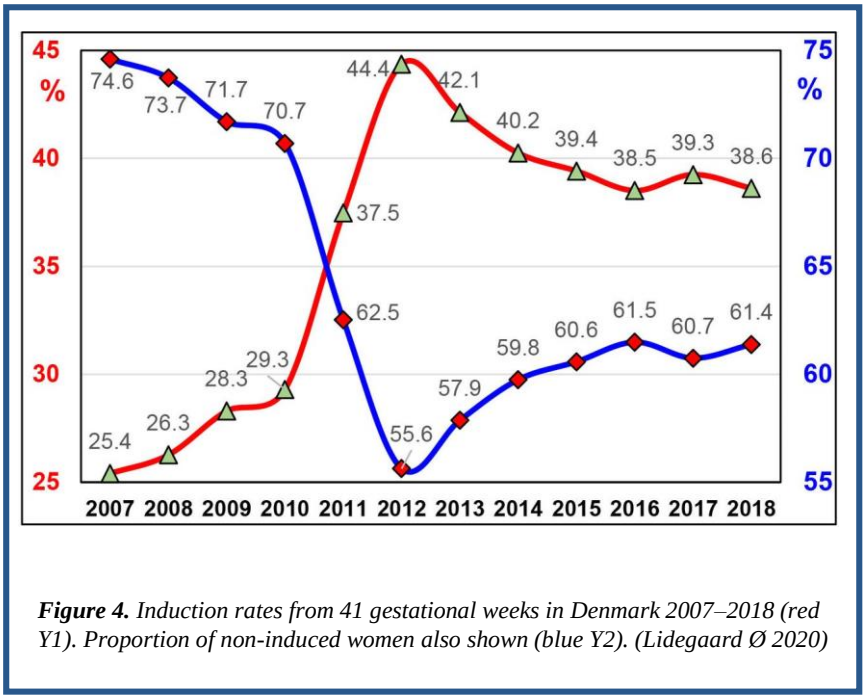

Figure 4. *Induction rates from 41 gestational weeks in Denmark 2007–2018 (red Y1). Proportion of non-induced women also shown (blue Y2). (Lidegaard Ø 2020)*

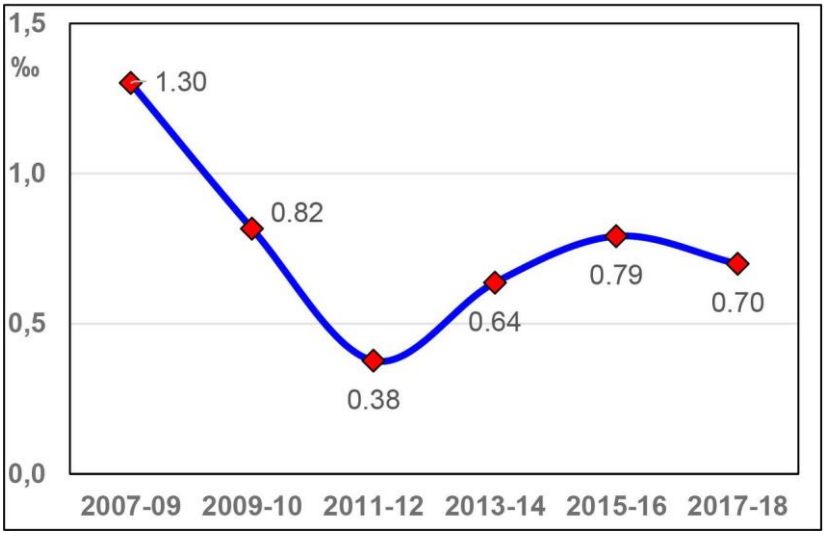

Figure 5. *Stillbirth rates per 1000 born from 41 gestational weeks from 2007 to 2018. (Lidegaard Ø 2020)*

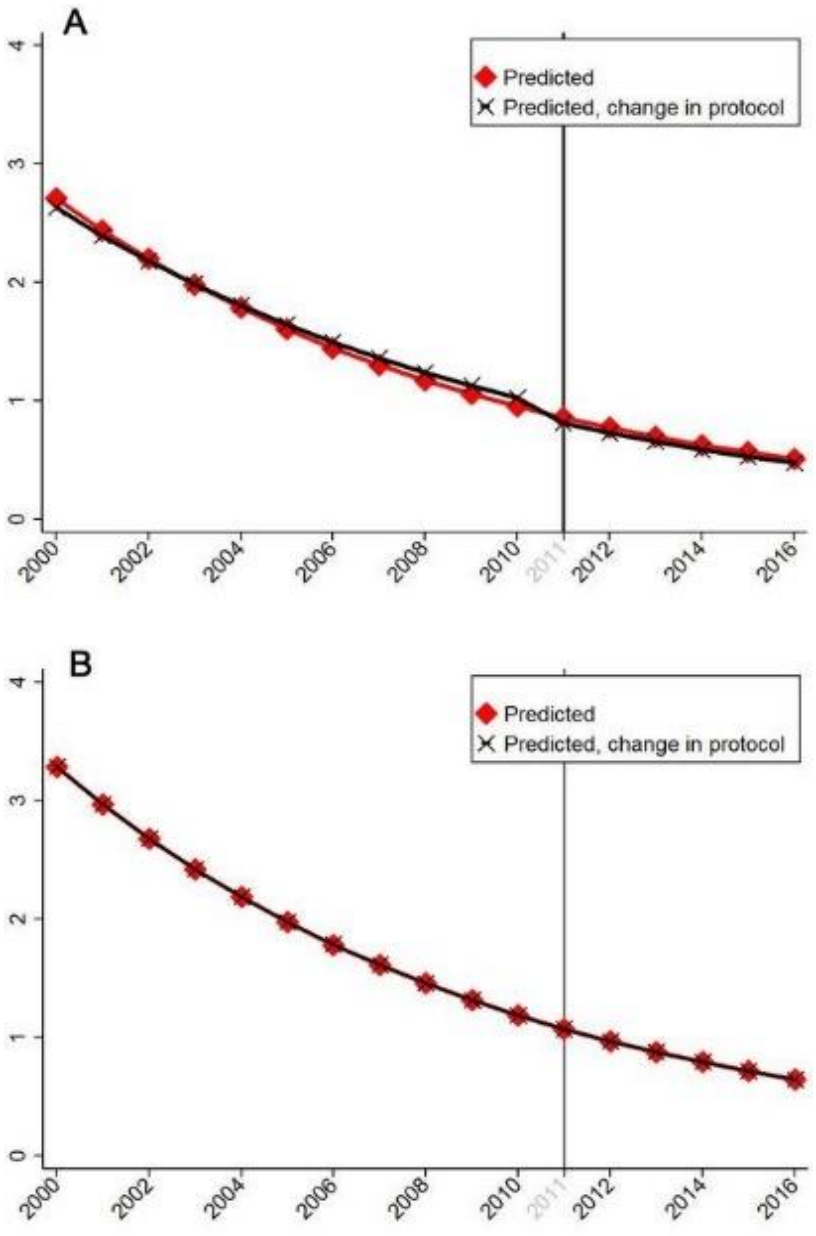

Figure 6. *Perinatal outcomes, year 2000–2016 with change in protocol, 2011. (A) Stillbirths per 1000 births (B) Perinatal death per 1000 births. (Rydahl, Declercq, et al., 2019)*

Health economic evaluations

In modern healthcare today it is important that any change of practice is fully assessed concerning both medical evidence as well as health economical aspects. The most common way of evaluating health economical aspects regarding a research question like ours is to perform a cost-effectiveness analysis (CEA). When performing a CEA, you compare relative costs and outcomes regarding an intervention and calculate an incremental cost effectiveness ratio (ICER). The ICER is calculated by dividing the difference in costs for the intervention group and control group with the difference in health gained or lost. A more detailed description is presented in Figure 8 in the section on statistical analysis under the subheading 'Paper III'. The most common measure of health gained or lost is quality adjusted life years (QALYs), including both a quality and quantity aspect of health. The quality component is measured on a scale ranging from 0 (dead) to 1 (perfect health) and the quantity consists of the time spent in the health condition. QALY is calculated by the multiplication of the quality and quantity measure. Life years (LYs) gained or lost is also frequently used. To perform a CEA a difference between intervention and control is needed. If no difference in health gained or lost is present a cost-minimization analysis can be performed, comparing the cost for the intervention group with the cost for the control group (Svensson, 2019).

Today, information on health economic effects of IOL is not well explored. In 1995, within one of the largest (n=3 407) RCTs comparing IOL at ≥41+0 GW with expectant management until 43+0 GW (Hannah et al., 1992) a CEA was planned. However, due to non-significant differences in perinatal outcomes between the groups a cost-minimization analysis was made instead (Goeree et al., 1995). It showed that IOL was associated with a lower cost as compared with expectant management of postterm pregnancies (Goeree et al., 1995). However, this study was made more than 25 years ago and might not be relevant today.

Another article in the field was published in 2011 assessing IOL at 41 GW versus expectant management with antenatal fetal surveillance (CTG and measurement of amniotic fluid index) in nulliparous women with a decision-analytic model (Kaimal et al., 2011). A decision analysis allows the analyst to compare the anticipated consequences of the different policies after taking into

account all relevant events and complications with their probabilities and considering all relevant clinical outcomes and costs. The ICER was $10 945 per QALY gained (the range of cost-effectiveness thresholds is commonly set at $50 000-100 000 per QALY in the USA). Thus, the conclusion was that it was cost effective to induce labour compared with expectant management (Kaimal et al., 2011). However, the calculations were based on a theoretical cohort with a caesarean delivery rate of 27 % in the induction group, which is much higher than e.g., in the INDEX trial (Keulen et al., 2019).

More recent studies from UK and US have been published (Einerson et al., 2020; Walker et al., 2017) and both are performed alongside a RCT comparing IOL with expectant management. Walker et al. conducted a CEA (n=380, 61 % of the RCT population) in UK of IOL at 39+0-39+6 GW in nulliparous women ≥35 years old compared with expectant management until 41+0 to 42+0 GW (Walker et al., 2016; Walker et al., 2017). They estimated a reduced mean cost of £263 with IOL, a mean difference in QALY in favour of IOL of 0.002 and an ICER of – £114 526. Hence, they concluded that IOL would save money (Walker et al., 2017). A cost-minimisation analysis (Einerson et al., 2020), from US was conducted on a subgroup (n=1 201, 20 % of the RCT population) of the ARRIVE trial (Grobman et al., 2018) (IOL at 39+0-39+4 GW in nulliparous women compared with expectant management and IOL at 41+0 GW). In that study they reported no differences in total costs for IOL compared with expectant management, but maternal intrapartum and delivery costs were higher in the IOL group and outpatient costs were lower. Neonatal care costs were similar in both groups (Einerson et al., 2020).

In summary, there is no recent study on cost-effectiveness regarding IOL at 41+0-2 GW compared with expectant management until 42+0-1 GW in a government funded healthcare system. Hence, more research regarding healthcare costs associated with IOL in low-risk women is needed.

Methods for induction of labour

In order to induce labour, in some cases you need to ripen the cervix in order to perform amniotomy and start the delivery. There are in principle two different methods for ripening the cervix; a mechanical and a pharmaceutical

approach. The mechanical method mainly consists of using a transvaginal balloon catheter (TVBC) inserted through the cervix with the inflatable part just above the internal os. In case of a double balloon, one balloon will be positioned below the external os. The balloons are inflated with 30–80 mL NaCl, Other mechanical methods, not as commonly used as the TVBC method, consists of membrane sweeping, nipple stimulation, extra-amniotic saline infusion using infusion rates of 0–40 mL/h, hygroscopic dilators or osmotic dilators (Laminaria japonicum). The pharmaceutical approach comprises of low-dose synthetic prostaglandin E1 (misoprostol), administered either vaginally or orally and synthetic prostaglandin E2 (dinoprostone) administered vaginally. Which approach to use, a mechanical or pharmacological, is mostly dependent on clinical tradition, but also partly dependent on how ripe the cervix is from start. Commonly the Bishop scoring system is used to describe the status of the cervix during pregnancy (Bishop, 1964; Edwards et al., 2000). If the Bishop's score is ≥six in nulliparous and ≥five in parous women the recommendation is to perform artificial rupture of the membranes, hence amniotomy, and start an oxytocin infusion in lack of contractions. A Bishop's score below five requires ripening of the cervix (ACOG, 2009; G. NICE, 2008; SFOG, 2016).

In October 2019, a Cochrane review, including 113 trials and 22,373 women, regarding mechanical methods for IOL, regardless of indication for IOL, was published (de Vaan et al., 2019). The aim of the review was to evaluate the efficacy and the safety comparing mechanical methods to low-dose misoprostol vaginally or orally, dinoprostone vaginally or intracervical and amniotomy or oxytocin. The conclusion was that there is moderate certainty of evidence, according to GRADE (G. H. Guyatt et al., 2008), that IOL with a TVBC versus low-dose misoprostol *orally* might increase the risk of a vaginal delivery not achieved within 24 hours (RR 1.28, 95 % CI 1.13 to 1.46) and slightly increase the caesarean delivery rates (RR 1.17, 95 % CI 1.04 to 1.32). Furthermore, it was unclear whether there was a difference or not regarding uterine hyper-stimulation with fetal heart rate changes, serious neonatal morbidity or perinatal death, serious maternal morbidity or death, five-minute Apgar scores < 7 and NICU admissions. Moreover, a TVBC seems to increase the risk of vaginal delivery not achieved within 24 hours compared with low-dose misoprostol *vaginally*. However, a TVBC compared with low-dose misoprostol vaginally might decrease the risk of uterine hyper-stimulation with

fetal heart rate changes (RR 0.39, 95 % CI 0.18 to 0.85), but increases the risk of caesarean delivery (RR 1.28, 95 % CI 1.02 to 1.60). As in the comparison with low-dose misoprostol orally it is also uncertain if there are any differences in serious neonatal morbidity or perinatal death, serious maternal morbidity or death, five-minute Apgar scores < 7 and NICU admissions. Lastly, comparing TVBC with vaginal dinoprostone, no difference in effectiveness was seen i.e. vaginal delivery not achieved within 24 hours (RR 1.01, 95 % CI 0.82 to 1.26) or rate of caesarean delivery (RR 1.00, 95 % CI 0.92 to 1.09). However, TVBC may decrease the risk of uterine hyper-stimulation with fetal heart rate changes (RR 0.35, 95 % CI 0.18 to 0.67), serious neonatal morbidity or perinatal death (RR 0.48, 95 % CI 0.25 to 0.93) and might reduce the risk of a NICU admission (RR 0.82, 95 % CI 0.65 to 1.04). The risk profile for serious maternal morbidity and mortality or five-minute Apgar score <7 is unclear (de Vaan et al., 2019). In the review only eight RCTs evaluated oral misoprostol (OM) compared with TVBC and of those only two small trials included late term pregnancies exclusively (IOL at 40+6 GW) (Goonewardene et al., 2014; Somirathne et al., 2017)

In summary, mechanical induction with a TVBC seems to be less effective than low-dose misoprostol orally and vaginally, but as effective as dinoprostone. It seems to be safer for the neonate with a TVBC induction than dinoprostone and vaginal low-dose misoprostol, but the safety profile remains uncertain for oral low-dose misoprostol versus a TBVC. Further research regarding efficacy and safety with low-dose misoprostol is needed specifically in late term/postterm pregnancies.

AIM OF THE THESIS

The aim of this thesis was to, in healthy women with a low-risk singleton pregnancy, clarify if a policy of IOL at 41+0-2 GW was superior, in terms of neonatal and maternal outcomes as well as health economic aspects as compared with expectant management and induction at 42+0-1 GW. Furthermore, different methods for cervical ripening and IOL in late term and postterm pregnancies were also assessed.

The specific aims were:

- To examine if IOL at 41+0-2 GW compared with expectant management and IOL at 42+0-1 GW was superior in terms of perinatal outcome in healthy women with a low-risk pregnancy.

- To examine if IOL at 41+0-2 GW compared with expectant management and IOL at 42+0-1 GW was superior in terms of perinatal and maternal outcomes overall and in different subgroups such as parity (nulliparous and multiparous), maternal age (<35 years and ≥35 years), and BMI (<30 and ≥30).

- To examine total cost per birth and cost-effectiveness, including costs for the woman and her child, of IOL at 41+0-2 GW compared with expectant management and IOL at 42+0-1 GW.

- To examine if there were any differences regarding efficacy, safety or women's experience between IOL in healthy women with a low-risk pregnancy and unripe cervix at 41+0 GW to 42+1 GW with the two preferred methods of cervical ripening; OM and TVBC.

PATIENTS AND METHODS

Setting and study design

This thesis consists of four studies with different study design, but all of them are based on or includes Paper I. The study design, settings and analysis for each study are summarised in Table 4.

Table 4. *Study design, setting and analysis*

	Paper I	**Paper II**	**Paper III**	**Paper IV**
Setting	14 centres in Sweden	Turkey, The Netherlands, Sweden	14 centres in Sweden	14 centres in Sweden
Study design	Register based superiority RCT	One-step IPD-MA	CEA	Prospective cohort study
Study period (randomisation)	2016-2018	Trials were published 2005 - 2019	2016-2018	2016-2018
Study Population	Low-risk 41+0 – 42+01 GW n=2 760	Low-risk 41+0 – 42+01 GW IPD: n=4 561 AD: n=5161	Low-risk 41+0 – 42+01 GW n=2 746	Low-risk OM n=744 TVBC n=469 (TVBC ref) n=1 213
Intervention group	IOL, n=1 381	IOL IPD, n=2 281 IOL AD, n= 2 581	IOL, n=1 373	NA
Control Group	EM, n=1 379	EM IPD, n=2 280 EM AD, n=2 580	EM, n=1 373	NA
Data Source	SPR, SNQ, SCB and eCRF	Data collected from included studies	Each participating centres' accountant department, SPR, SNQ, SCB and eCRF	SPR, SNQ, SCB and eCRF
Primary Outcome	A composite of PNM and morbidity	A composite of PNM and morbidity	Cost/LY and cost/QALY	*Efficacy*: vaginal delivery within 24 hours *Safety neonatal*: composite of PNM and morbidity *Safety maternal*: composite of mortality and morbidity
Main Analysis	IOL versus EM*	IOL vs EM*	IOL vs EM*	OM vs TVBC
Subgroup Analysis	Maternal age (<35 and ≥35 years), parity (nulliparous and parous), BMI (<30 and ≥30)	Maternal age (<35 and ≥35 years), parity (nulliparous and parous), BMI (<30 and ≥30) and fetal sex (male and female)	Robustness analysis on the Gbg cohort using a more precise cost data (cost/hour instead of cost/day)	Parity (nulliparous and parous) and BS (0-4 and 5-10)

AD, aggregate data; BMI, body mass index; BS, Bishop score; CEA, cost-effectiveness analysis; eCRF, electronic case report form; EM, expectant management; Gbg, Gothenburg; IOL, induction of labour; IPD, individual participant data; LY, life year; MA, meta-analysis; OM, oral misoprostol; Ref, reference; SCB, Statistics Sweden; SNQ, Swedish Neonatal Quality register, PNM, perinatal mortality; SPR, Swedish Pregnancy Register; TVBC, transvaginal balloon catheter; QALY, quality adjusted life years; vs, versus
*On the intention to treat population

Paper I is a Swedish multicentre register-based randomised controlled superiority trial in healthy women with a low-risk late-term/postterm singleton pregnancy and a fetus in cephalic presentation. The trial was conducted in 14 centres, five university hospitals and nine county hospital, between 2016 and October 2018. Paper II is a one-step individual participant data (IPD) MA including trials comparing IOL in women with a low-risk singleton pregnancy and a fetus in cephalic presentation at 41+0-2 GW with expectant management until 42+0-1 GW. Paper III is a CEA alongside Paper I. Paper IV is a prospective cohort study investigating the efficacy, safety and women's experiences of IOL with OM compared with TVBC. The population studied was women included in Paper I in need of cervical ripening for IOL.

Data sources

Paper I

In Paper I, data were collected from the SPR, Swedish Neonatal Quality register (SNQ), Statistics Sweden and from individual electronic medical record (EMR). Information from individual EMR was recorded in an electronic case report form (eCRF). The unique personal identification number assigned to all persons born or immigrated to Sweden was used to link eCRF data with data from the three registers.

The National Board of Health and Welfare administrates a number of national health data registers, e.g. the Swedish Medical Birth Register (MBR), to facilitate analyses and development of Swedish healthcare and social services. These health data registers are regulated by Swedish law and they are subject to strict confidentiality. Healthcare services are obliged to report to the registers. SPR and SNQ do not have health data register status, but are national quality registers, certified by The Swedish Association of Local Authorities and Regions (SALAR) and administrated by the profession. Hence, it is not mandatory to contribute with data. However, the coverage of healthcare providers regarding SPR is >90 % and for SNQ it is 100 %. The rate of individual patients declining participation in the registers is <1 % (M. Norman et al., 2019; Stephansson et al., 2018). We chose to work with the quality

registers (SPR and SNQ) and not the Swedish MBR since the former are more accessible and can provide more detailed information.

The Swedish Pregnancy Register (SPR)
The SPR is a national quality register and was initiated by Swedish healthcare professionals in 2013. It was a merge of two pre-existing registers, the Maternal Healthcare Register (started in 1999) and the National Quality Register for Prenatal Diagnosis (started in 2006). The register collects data with start at the first antenatal care visit, during the whole pregnancy and delivery and the data collection stops at the follow-up visit to the antenatal care unit. Hence, the data are collected from about 5-8 GW to 16 weeks postpartum. Data collected concerns demographics, reproductive and maternal health information, information on prenatal diagnostics e.g., ultrasound examinations, delivery outcomes and perinatal and maternal outcomes. Data are collected from three different sources:

1. Automatic transfer of approximately 220 variables from EMR. The variables include antenatal care information, biometry from ultrasound examinations performed during pregnancy, data on delivery outcome and postpartum care including international classification of disease (ICD) codes, surgical codes and codes for other interventions e.g., CTG, intravenous infusions and intravenous administration of drugs.

2. Data that is not registered in the EMR is manually entered via the web by midwives at the first antenatal visit and the follow-up visit eight to 16 weeks postpartum. It includes data, such as, country of birth, level of education, main occupation, self-rated health before pregnancy, prenatal diagnosis, use of professional translator, parent support attendance, support for fear of childbirth, treatment of psychiatric disorders, screening for intimate partner violence, oral glucose test values, physician attendance in antenatal care, maternal weight postpartum, breast feeding at four weeks postpartum and self-reported health during and after pregnancy.

3. Data automatically transferred from a web-based system for calculating the likelihood ratio for trisomy 13, 18 and 21 based on a first trimester combined ultrasound and biochemistry examination.

In Sweden, between 2016 and 2018, three different EMRs were in use and 90 % of the antenatal and delivery care used the EMR Obstetrix©. From Obstetrix© there was a transfer of data 24 hours after birth. The two other EMRs were not fully connected to SPR at the time of our first study. However, only one centre did not use Obstetrix© in our study, Uppsala University Hospital, and data not transferred from their EMR ("Cosmic") automatically was manually entered in a separate data file. Both internal and external validation of the registers' quality have been made and the data entered via the web by midwives showed good (70-94 %) or very good (≥95 %) agreement with medical records (Petersson et al., 2014). The external validation was made with the MBR for deliveries in 2015 and the coverage was 98-100 %, when direct transfer of data from the EMRs was done (Stephansson et al., 2018).

Swedish Neonatal Quality Register (SNQ)

In 2001, SNQ was launched and all neonatal units in Sweden (n=37) are connected to the register. SNQ collects data on over 300 variables regarding newborns admitted to neonatal units during the first four postnatal weeks. Information on 27 medical, four surgical and three organisational key performance indicators in neonatal care is collected. Additionally, 25 outcomes including neonatal mortality and morbidity and health status at follow-up visits are listed in SNQ. Furthermore, parental reported experience measures are included as well as ICD codes, surgical codes and codes for other interventions. Maternal, pregnancy and delivery data from EMRs are automatically transferred via the SPR to SNQ. Data regarding the infant were, during the study period, manually extracted and logged into a web report. Data from neonatal transport teams are automatically logged into SNQ. The SNQ register has excellent completeness and high validity (M. Norman et al., 2019).

Statistics Sweden

The authority responsible for official Swedish statistics is called Statistics Sweden. Their assignment is to complete and present official statistics regarding 22 subject areas and 112 statistical areas including healthcare information on cause of death, health and disease (on aggregate level),

population changes, information on income, socio-economic conditions, educational level and country of birth on individual level. In order to gather data related to healthcare a collaboration between Statistic Sweden and the Swedish National Board of Health and Welfare, a government agency under the Ministry of Health and Social Affairs, is vital. This agency is in turn responsible for health data registers such as the Swedish MBR, the National Patient Register and Cause of Death Register. In our study, we utilized statistics from Statistics Sweden regarding vital status regarding mother and infant.

Electronic case record form

Data that were not present in any of the three registers mentioned above were extracted from EMR and manually entered into the eCRF after the mother and infant were discharged from the delivery stay. Variables included in the eCRF were randomisation group, eligibility for the trial, antenatal visits from randomisation until delivery and stay in hospital from randomisation until delivery. In addition, IOL according to protocol or not, including reason for deviation from the protocol, IOL methods used and Bishop score assessments during the IOL process were also recorded in the eCRF. Additionally, time point for caesarean or operative vaginal delivery during the birth process, signs of infection during labour, use of antibiotics as prophylaxis or therapy during labour and admission to intensive care unit were entered into the eCRF. Furthermore, a module in the eCRF was used for the randomisation process.

Paper II

In Paper II IPD were sought by contacting corresponding authors of the included trials. The Swedish postterm induction study (SWEPIS) (Wennerholm et al., 2019) and the INDEX trial (Keulen et al., 2019) contributed with IPD. The Turkish trial (Gelisen et al., 2005) declined to participate with IPD, due to difficulties retrieving the data. Aggregated data from the Turkish trial were used for perinatal mortality and rate of caesarean delivery.

Paper III

Data sources in Paper III included outcome data from SWEPIS regarding perinatal, maternal and delivery outcomes (Wennerholm et al., 2019). Cost per patient (CPP) data were collected from the accountancy department at each included centre in SWEPIS.

Cost per patient (CPP)

SALAR has developed a system called CPP in order for different healthcare providers to easily benchmark costs (SALAR, 2020). The CPP system consists of fixed and variable costs and reports costs per care event. Regarding hospital stay, it reports a standardised fixed cost per day and not per hour. The fixed costs include costs for staffing, rent, everyday materials, depreciation and everyday laboratory costs. Variable costs include e.g., costs for being in the operating theatre and/or intensive care unit, additional costs for advanced laboratory work, and imaging diagnostics.

Quality of life assessment

Health-related quality of life effects for the women were estimated by means of the EuroQol -5 Dimension measure (EQ-5D) QALY calculations (Rabin et al., 2001). EQ-5D assesses mobility, self-care, usual activity, pain/discomfort and anxiety/depression. The EQ-5D was assessed in a subpopulation of women participating in SWEPIS at three centres representing one large delivery unit, Sahlgrenska University hospital with 10 000 deliveries per year, one university hospital, Örebro University Hospital with 3 000 deliveries per year and one county hospital, Falu Hospital with 3 200 deliveries per year. Women filled out the questionnaire at randomisation and 3 months after delivery. LYs were measured based on differences in perinatal mortality between groups in Paper I multiplied by Swedish survival probabilities from life table statistics (Statistics Sweden, 2020).

Paper IV

In Paper IV the same data sources as in Paper I were used. In addition, information regarding women's childbirth experience was collected using the Childbirth Experience Questionnaire (CEQ 2.0) and Visual Analogue Scale (VAS) (Dencker et al., 2020; Dencker et al., 2010; Register, 2019).

Childbirth Experience Questionnaire (CEQ 2.0)

The CEQ 2.0 is a validated questionnaire assessing women's childbirth experience by exploring four different domains of interest. The domains are own capacity (eight items), perceived safety (six items), professional support (five items) and participation (three items) (Figure 7).

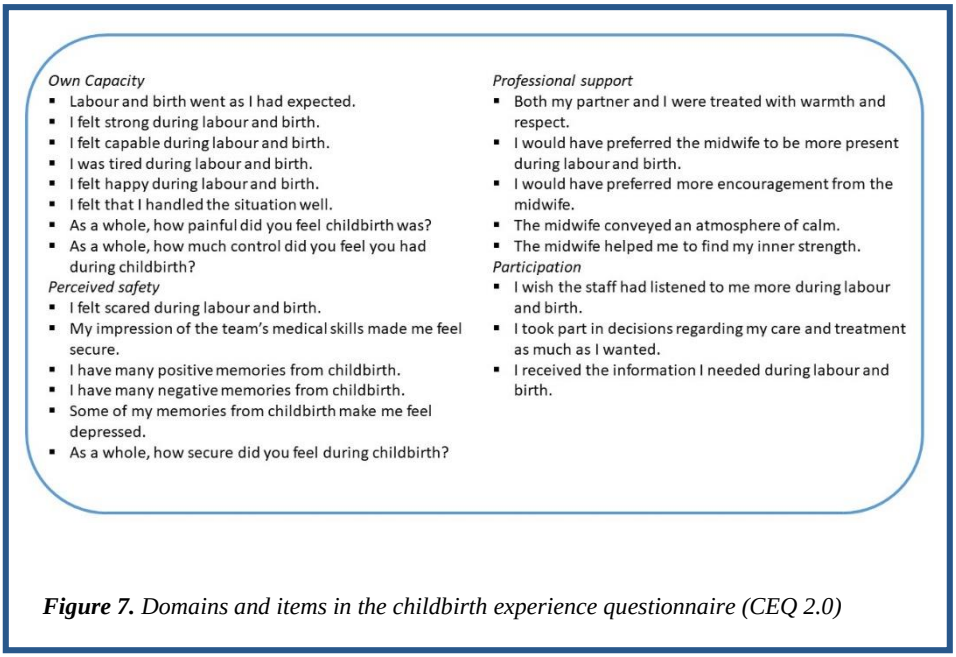

Figure 7. *Domains and items in the childbirth experience questionnaire (CEQ 2.0)*

Items are evaluated on a 4-point Likert scale, ranging from 1=totally disagree, 2=mostly disagree, 3=mostly agree and 4=totally agree. The items regarding pain, control and sense of security are rated with a VAS 1-100, which are categorized as 0-40=1, 41-60=2, 61-80=3 and 81-100=4. Any negatively worded items are reversed in scoring. The higher scores the more positive childbirth experience. The total CEQ 2.0 score is the mean of the four individual domain scores.

The questionnaire was retrieved from the same subgroup of women answering the EQ-5D questionnaire, thus, from a subgroup of women representing different geographical areas in Sweden and different sizes of delivery units.

Visual Analogue Scale (VAS)
VAS was used to estimate women's overall childbirth experience within three days after delivery. It was retrieved from women in the hospitals where CEQ 2.0 was not distributed. The VAS ranged from 1, representing a very negative experience, to 10, representing a very positive experience (Register, 2019). Rating 8-10 was considered to be a very good childbirth experience and ratings of 1-2 as very bad.

Definitions of outcome variables

Relevant outcome measures regarding management of late term pregnancy includes perinatal adverse outcomes, maternal adverse outcomes, delivery/efficacy outcomes, women's experiences and healthcare cost outcomes. Perinatal and maternal mortality and most severe neonatal and maternal morbidity outcomes are rare and therefore difficult to use as individual outcomes in both RCTs and cohort studies, as illustrated by the Cochrane report from 2005 (Pattinson et al., 2005). Hence, it is not unusual to combine perinatal mortality and morbidity into a composite outcome indicator (Chen et al., 2019; Crowther et al., 2006; Keulen et al., 2019; Tita et al., 2011).

Aspects influencing the decision of components included in a composite outcome in a study are the relationship to the intervention or the control treatment. In a recommendation on core outcomes related to IOL published in 2018 they propose, in addition to perinatal mortality, indicators of birth asphyxia (hypoxic ischaemic encephalopathy (HIE), MAS, neonatal convulsions and need for respiratory support), birth trauma, neonatal infection and admission to NICU (Dos Santos et al., 2018). In the same recommendation core maternal outcomes are presented including cardiorespiratory arrest, damage to internal organs (bowel, bladder or ureters), haemorrhage, hysterectomy for any complications resulting from birth, intensive care unit admission, length of hospital stay, maternal death, maternal infection, maternal satisfaction, mode of delivery, more than one induction agent required, oxytocin augmentation, postnatal depression, pulmonary embolus, stroke, time from IOL to delivery, uterine hyperstimulation and uterine scar dehiscence or rupture (Dos Santos et al., 2018).

We decided to use an adverse perinatal composite indicator outcome as primary outcome in Paper I, II and IV. Paper I was planned before Dos Santos

et al. (Dos Santos et al., 2018) published their core outcome recommendations for studies on IOL. However, despite this we included most of the recommended outcomes except hyperstimulation of the uterus. We decided to include, in our primary adverse perinatal composite indicator outcome, perinatal mortality, indicators of asphyxia such as Apgar score <7 at five minutes, metabolic acidosis, HIE I-III, neonatal convulsions, MAS and mechanical ventilation within the first 72 hours. Furthermore, signs of birth trauma such as brachial plexus injury and intracranial haemorrhage were included. The components of the primary perinatal adverse composite indicator differed slightly between the three papers as described below.

Adverse perinatal composite indicator outcome

Perinatal mortality

Perinatal mortality is defined as stillbirth and early neonatal death. WHO/ICD defines stillbirth as the death of a fetus during pregnancy or labour that has reached a birth weight of 500 g, or if birth weight is unavailable, gestational age of ≥22 weeks or crown-to-heel length of ≥25 cm (Lawn et al., 2010). Early neonatal death is defined as death of a live-born infant during the first seven completed days after birth (day 0-6) and late neonatal death is defined as death of a live-born infant during the first 28 completed days after birth (day 0-27) (WHO, 1977). We excluded deaths due to accidents or lethal malformation not known before randomisation.

Apgar score

Apgar score is a swift approach to assess clinical status of a newborn during the first 10 minutes of life (Apgar, 1953). It was first introduced by Virginia Apgar in 1953 and comprises five assessment criteria; skin colour, pulse rate, reflex irritability grimace, muscle tone and respiratory effort. Each criterion can get a minimum of zero points and a maximum of two points. Hence, the maximum score indicating the status of an infant as good is 10.

Apgar score is not a predictor of neurologic outcome. However, the American College of Obstetricians and Gynaecologists (ACOG) states a five minute Apgar score of 7-10 as reassuring, a score of 4-6 as moderately abnormal and 0-3 as low in the term infant or “late-preterm infant” (ACOG, 2014). An

outcome with Apgar score <7 at five minutes has been regarded as the golden standard as a predictor of neonatal morbidity and used in many publications on neonatal adverse outcomes. However, after ACOG published a committee opinion in 2015 the new golden standard is regarded as Apgar <4 at five minutes (ACOG, 2015). The rationale for the change in recommendation is that Apgar score of 0-3 at five minutes correlate better with neonatal mortality and morbidity in large populations than Apgar <7(Casey et al., 2001; Li et al., 2013; Nelson et al., 1984).

Metabolic acidosis

Different definitions of metabolic acidosis in umbilical cord artery in newborns have been used. According to ACOG and the American Academy of Pediatrics metabolic acidosis in the newborn is defined as pH < 7.00 or base deficit ≥ 12 mmol/L (or > 16 mmol/L for therapeutic hypothermia) or both (ACOG, 2014). However, the definition pH<7.05 and base deficit >12 mmol/l has been widely used (Amer-Wåhlin et al., 2001; Wiberg-Itzel et al., 2008). We choose to use this definition or pH<7.00 (with or without base deficit) (Wiberg-Itzel et al., 2008). The reason for using pH<7.00 (with or without base deficit) was due to the difficulties in retrieving a blood sample where both pH and base deficit were present. The definition of a valid blood sample was if pH in the umbilical artery was lower than in the umbilical vein and if the arterial partial pressure of carbon dioxide was greater than the pressure in the vein. A sample was also regarded as valid if pH in the umbilical artery was below 7.00 (no venous blood sample needed).

Hypoxic ischaemic encephalopathy (HIE) I-III

HIE is the result of a hypoxic ischemic event affecting the brain and a clinical diagnosis based on neonatal encephalopathy in combination with low Apgar score and/or metabolic acidosis in the umbilical cord artery (Volpe, 2012). HIE is divided into three stages, mild (I), moderate (II) and severe (III) and each stage has its own ICD code. The diagnostic criteria used in Sweden is based on the classification by Sarnat (Sarnat et al., 1976).

Neonatal convulsions

Our definition of neonatal convulsions was based on the ICD-10 diagnostic criteria and included convulsions with or without electroencephalography

confirmation. Electroencephalographic activity compatible with convulsions without clinical signs of convulsions was also included in the definition. Most common were clinical convulsions i.e., twitching either in a body part, or throughout the body, or e.g., hiccups or similar. It is important to distinguish this from benign sleep myoclonus, or other benign twitching. If electroencephalography was completely normal despite suspected clinical convulsions, it was a clinical decision whether the infant was diagnosed with neonatal convulsions or not.

Meconium aspiration syndrome (MAS)
In Sweden, MAS is defined as a syndrome including meconiumstained amniotic fluid in combination with severe respiratory distress with need for mechanical ventilation after birth. A chest X-ray displaying asymmetric patchy pulmonary opacities is usually present, however not mandatory.

Mechanical ventilation within the first 72 hours
The definition of this outcome was that a tracheal tube was inserted.

Intracranial haemorrhage
The diagnosis intracranial haemorrhage was based on imaging diagnostics such as ultrasound, computed tomography or magnetic resonance imaging.

Brachial plexus injury
Brachial plexus injury is a clinical diagnosis and includes Erb Duchennes palsy (level C5 – C7 are affected in the spinal cord), Déjèrine-Klumpkes palsy, (level C7 – Th1 are affected in the spinal cord) and total damage (level C5 – Th1 are affected in the spinal cord) (Dodds et al., 2000).

Paper I

The primary adverse perinatal composite indicator outcome was a combination of perinatal mortality and neonatal morbidity as described in the previous section.

Secondary neonatal outcomes included admission to NICU and indications, such as e.g., neonatal jaundice, neonatal infections, hypoglycaemia and birth

trauma. In addition, birthweight related outcomes such as small for gestational age and macrosomia were also presented.

Secondary delivery outcomes included e.g., mode of delivery and duration of labour and shoulder dystocia. Secondary maternal outcomes including e.g., hypertensive disorders of pregnancy, perinatal lacerations III and IV postpartum haemorrhage (>1000 ml), infections and mortality within 42 days after birth.

Paper II

In Paper II the primary adverse perinatal composite indicator outcome was adjusted according to the two trials that contributed to the IPD-MA. Definitions of variables were compared and synchronised and a new primary adverse perinatal composite indicator outcome was agreed upon. Changes made compared to the primary outcome in Paper I were Apgar <4 instead of <7 at five minutes, HIE II-III instead of HIE I-III and removal of metabolic acidosis. Additional perinatal outcomes in Paper II included e.g., admission to neonatal medium or intensive neonatal care, admission to neonatal care ≥ 4 days as an indicator of sick newborns in need of longer treatment and/or more extensive observation and, as in Paper I, birthweight related outcomes such as small for gestational age and macrosomia.

Secondary delivery outcomes such as e.g., oxytocin during labour, pain treatment during vaginal delivery and mode of delivery were reported. Secondary maternal outcomes included e.g., hypertensive disorders of pregnancy, perineal lacerations III and IV, postpartum haemorrhage, fever during labour (≥38˚C), intravenous antibiotics during labour (prophylaxis or therapy) and maternal death up to 42 days after delivery.

Risk of bias

In Paper II all included trials were assessed using the risk of bias tool developed by Cochrane (Higgins et al., 2011). Risk of biases were adequacy of randomisation (selection bias), blinding for participants and personnel and statistician responsible for analysis (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective

reporting (reporting bias), and other bias (conflict of interest). The risks were reported as low, moderate, high or unclear.

Paper III

Primary outcome in Paper III was defined as ICER per LY and per QALY. Total cost of birth included costs associated with the delivery from admittance to discharge, costs for outpatient visits and inpatient stay between randomisation and admittance to the delivery ward and costs for neonatal care (in association with the delivery, inpatient stay and/or home care).

Secondary outcomes were the individual components of the total birth cost individually.

Paper IV

Paper IV has three primary outcomes including efficacy, neonatal safety and maternal safety.

Efficacy outcome
The efficacy outcome was chosen in order to reflect both efficacy in time and mode of delivery and consisted of vaginal delivery achieved within 24 hours from start of induction.

Neonatal safety outcome
The primary safety outcome was defined as a neonatal adverse composite indicator outcome, including intrapartum and neonatal mortality (stillbirths that occurred before start of IOL and deaths due to accidents or lethal malformation not known before randomisation were excluded) and morbidity. Components of morbidity were chosen according to the core outcomes presented in the article of Dos Santos et al. from 2018.

Maternal safety outcome
The maternal safety outcome was defined as an adverse composite indicator outcome including maternal mortality and morbidity. The maternal morbidity

components were also chosen according to core outcomes presented in Dos Santos et al., 2018.

Women's childbirth experience
CEQ 2.0 and VAS were used to measure women's childbirth experience.

Statistical analysis

All statistical methods are presented in detail in each paper. An overview of the statistical methods is presented in Table 5. Significance level was set to 0.05. All analyses were performed with SAS System Version 9 for Windows (SAS, Cary, NC).

Paper I

The statistical analysis in Paper I was made in collaboration with professional statisticians at Statistical Consulting Group, Gothenburg, Sweden. A summary of the statistical analyses is presented in Table 5. The expectant management group was the reference group and the primary analysis was made on the intention to treat group. Analysis was also made in the per protocol population and pre-specified subgroup analyses were made on maternal age (<35 versus ≥ 35 years), parity (nulliparous versus parous women) and BMI (<30 versus ≥ 30). Whether the effect of treatment differed between subgroups was tested with logistic regression and the interaction term treatment x subgroup variable.

Paper II

In Paper II the statistical analysis was made in collaboration with one professional statistician at Statistical Consulting Group, Gothenburg, Sweden and one at Amsterdam UMC, University of Amsterdam, Department of Obstetrics and Gynaecology, Amsterdam Reproduction & Development Research Institute, Amsterdam, the Netherlands. A summary of the statistical analysis is presented in Table 5. The expectant management group was the reference group and the primary analysis was made on the intention to treat group. Peto odds ratio, in a two-step approach, was used when zero events in

one arm of one or both trials were present. If zero events in both arms in all but one trial, risk estimate and inferential statistics were not calculated due to double zero events will add zero weight to the IPD-MA.

Table 5. *Statistical methods used in the Papers*

	Paper I	**Paper II**	**Paper III**	**Paper IV**
Descriptive statistics	*Categorical variables:* number and percentage *Continuous variables:* mean (SD), median (IQR)	*Categorical variables:* number and percentage *Continuous variables:* mean (SD), median (IQR)	*Categorical variables:* number and percentage *Continuous variables:* mean (SD), median (IQR)	*Categorical variables:* number and percentage *Continuous variables:* mean (SD), median (max/min/IQR)
Comparison	IOL vs EM	IOL vs EM	IOL vs EM	OM vs TVBC
Main analysis population	ITT	ITT	ITT	
Analytical statistics	*Primary outcome:* Fisher's exact test (one sided p-value x 2) *Dichotomous variables:* Fisher's exact test and RR (95 % CI) *Continuous variables:* Fisher's non-parametric permutation test and MD (95 % CI) *Ordered categorical variables:* Mantel Haenszel χ^2 test *Non-ordered categorical variables:* Pearson's χ^2 test	**IPD**: One-stage MA. *Dichotomous variables*: RR and RD (95 % CI) using general linear models with categorical coding for study If zero events in one arm in any or both trials Peto OR was used. *Continuous variables:* MD (95 % CI) using a general linear model with categorical coding for study *Non-ordered categorical variables:* Pearson's χ^2 test *Heterogeneity:* I^2 and p-value of variability **AD:** RR (95 % CI) using Mantel-Haenszel fixed effect model or Peto OR	*Arithmetic means:* Linear regression where standard error and CI were based on non-parametric bootstrapping *ICER:* ratio difference in arithmetic mean costs and difference in arithmetic mean *LYs and QALYs:* CI for ICER was calculated with Fieller's theorem	*Categorical variables:* crude and adj RR and RD (95 % CI) were calculated with regression adj for PS *Continuous variables*: Adj means with SEM and mean differences were calculated with regression, adj for PS *Dichotomous variables:* Fisher's exact test *Continuous variables:* Fisher's non-parametric permutation *Non-ordered categorical variables:* Pearson's χ^2 test

AD, aggregate data; Adj, adjusted; CI, confidence interval: EM, expectant management; ICER, incremental cost-effectiveness ratio; IOL, induction of labour; IQR, interquartile range; IPD, individual participant data; ITT, intention to treat, MD, mean difference; LY, life year; OR, odds ratio; PS, propensity score; RD, risk difference; RR relative risk; QALY, quality adjusted life year; vs, versus

Pre-specified subgroup analysis on the primary outcome, perinatal mortality and caesarean delivery was made on maternal age (<30 years versus ≥35 years), parity (nulliparous versus parous women) and BMI (< 30 versus ≥30). In addition, a post hoc subgroup analysis was conducted on fetal sex. Statistical method used for the subgroup analysis was the same as in Paper I.

Paper III

Statistical analysis was performed by the senior author of the paper. A summary of the statistical analysis is presented in Table 5. The expectant management group was the reference and the primary analysis was made on the intention to treat group. ICER was calculated based on the difference in arithmetic mean in total cost per birth in the IOL group and the expectant management group divided by the difference in arithmetic mean in LYs as well as QALYs (Figure 8).

$$ICER = \frac{\text{Total birth costs}_{IOL} - \text{Total birth costs}_{EM}}{\text{Life years}_{IOL} - \text{Life years}_{EM}}$$

$$ICER = \frac{\text{Total birth costs}_{IOL} - \text{Total birth costs}_{EM}}{\text{QALY}_{IOL} - \text{QALY}_{EM}}$$

Figure 8. *Incremental cost-effectiveness ratio Paper IV*

Paper IV

The statistical analysis in Paper IV was made in collaboration with a professional statistician at Statistical Consulting Group, Gothenburg, Sweden. A summary of the statistical analysis is presented in Table 5. The TVBC group was used as reference. We used a propensity score (PS) model; PS as a covariate in a regression calculation, to control for confounders that might influence group assignment and outcome (Rosenbaum PR, 1983). PS was based on centre (n=15, Table S1 in Paper IV) and all baseline characteristics presented in Table 1 in Paper IV except for medical history of pre-gestational diabetes, medical history of chronic hypertension, height and last recorded weight during pregnancy. Crude and adjusted RR with 95 % CI were calculated and adjustments were made for PS.

In addition, Kaplan-Meier analysis and logrank tests on time-to-vaginal delivery and time-to-delivery was performed. The time-to-delivery analysis was also stratified by Bishop score and parity. Subgroup analysis on the three primary outcomes was made on Bishop score and parity.

Ethic approval

The Regional Ethics Board in Gothenburg approved the study in May 2014 (Dnr: 285-14) and later its complementary applications (T 905-15, T 291-16, T 1180-16, T 330-17, T 1066-17, T 087-18, T 347-18, T 961-18, T 1110-18, and INDEX-SWEPIS 2019-04094).

RESULTS AND COMMENTS

Paper I

Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (SWEdish Postterm Induction Study, SWEPIS): multicentre, open label, randomised superiority trial

A total of 10 038 women were planned to participate in this RCT, but the trial was stopped early due to safety reasons when 2 760 (1 381 in the IOL group and 1 379 in the expectant management group) women were included. The reason for early stopping of the trial was a significantly higher perinatal mortality rate in the expectant management group (p-value 0.03).

In the IOL group 85.5 % (1 181/1 381) of the women were induced and 14.1 % (195/1 381) went into spontaneous labour while awaiting IOL, 0.4 % (5/1 381) had a scheduled caesarean delivery. In the expectant management group 33.1 % (457/1 379) had IOL and 66.7 % (920/1 379) went into spontaneous labour before 42+0-1 GW, 0.1 % (2/1 379) had a scheduled caesarean delivery.

Perinatal outcomes

The primary perinatal adverse composite indicator outcome and statistically significant perinatal outcomes are presented in Table 6. The primary adverse perinatal composite indicator outcome did not differ between the groups. Perinatal mortality differed significantly with zero deaths in the IOL group and six in the expectant management group. There were five stillbirths and one neonatal death. The individual morbidity components separately or combined were similar in both groups. Furthermore, admission to NICU was significantly lower in the IOL group compared to the expectant management group.

Maternal outcomes

There were significantly fewer hypertensive disorders of pregnancy in the IOL group and significantly more women with the diagnosis of endometritis in the expectant management group (Table 7). Other outcomes such as postpartum haemorrhage, cervical lacerations, uterine rupture, perineal lacerations III and IV, venous thromboembolism and birth related infections were similar in both groups. There were no maternal deaths.

Table 6. *Summary of primary adverse perinatal composite indicator outcome and statistically significant perinatal outcomes*

Variables	Induction of labour (n=1 381)	Expectant management (n=1 379)	Relative risk (95 % CI)	p-value
Primary composite outcome*	33 (2.4)	31 (2.2)	1.06 (0.65 to 1.73)	0.90
PNM†	0 (0.0)	6 (0.4)	NA	0.03
Stillbirth	0 (0.0)	5 (0.4)	NA	0.06
Neonatal morbidity‡	33 (2.4)	26 (1.9)	1.27 (0.76 to 2,11)	0.43
Admission to NICU	55 (4.0)	82/1 374 (6.0)	0.67 (0.48 to 0.93)	0.02
Birth weight (g) Mean (SD)	3 815 (409)	3 875 (436)		<0.001
Small for gestational age§	9 (0.7)	22 (1.2)	0.41 (0.19 to 0.88)	0.03
Macrosomia (≥4 500g)	68 (4.9)	114 (8.3)	0.60 (0.45 to 0.80)	<0.001
Jaundice¶	16 (1.2)	32/1 374 (2.3)	0.50 (0.27 to 0.90)	0.03

Values are numbers (percentages) unless stated otherwise
CI, confidence interval; NA, not applicable; NICU, neonatal intensive care unit; PNM, perinatal mortality; SD, standard deviation
*Stillbirth, neonatal mortality, Apgar score ≤7 at five minutes, metabolic acidosis, hypoxic ischemic encephalopathy I-III, intracranial haemorrhage, neonatal convulsions, meconium aspiration syndrome, mechanical ventilation within first 72 hours, obstetric brachial plexus injury
†Stillbirth and neonatal mortality (live births with mortality <28 days)
‡Apgar score ≤7 at five minutes, metabolic acidosis, hypoxic ischemic encephalopathy I-III, intracranial haemorrhage, neonatal convulsions, meconium aspiration syndrome, mechanical ventilation within first 72 hours, obstetric brachial plexus injury
§Small for gestational age defined as ≤2 standard deviations, according to Swedish sex and gestational age specific reference (Marsal et al., 1996)
¶requiring phototherapy or exchange transfusion

Delivery outcomes

As shown in Table 7 there were significantly more use of epidural anaesthesia and less women with meconium stained amniotic fluid in the IOL group compared to the expectant management group. In addition, women in the IOL group had a longer hospital stay from admission to delivery, but a shorter duration of labour than the expectant management group. The duration of postpartum stay did not differ (IOL group; n=1 333, mean 46.3 hours [SD:

27.0] and expectant management group; n=1 333, mean 47.1 hours [SD: 29.7], mean difference between groups 0.46; 95 % CI −0.82 [−2.99 to 1.32]).

No difference in caesarean delivery rate between the groups was shown. Neither were there any differences in indication for caesarean delivery. Failure to progress was the leading cause with 51.4 % (71/138) in the IOL group and 52.7 % (77/146) in the expectant management group.

Table 7. *Mode of delivery and statistically significant delivery and maternal outcomes*

Variables	Induction of labour (n=1 381)	Expectant management (n=1 379)	Mean difference (95 % CI)/ Relative risk (95 % CI)	p-value
Time from admission to labour ward to delivery (hours) **Mean (SD)** **Median (IQR)**	n=1 390 20.1 (14.8) 16.2 (9.2-27.9)	n=1 378 13.6 (12.2) 10.4 (4.6-19.0)	6.49 (5.50 to 7.50)*	<0.001
Duration of labour† (hours) **Mean (SD)** **Median (IQR)**	n=717 7.13 (5.39) 5.67 (2.85-10.28)	n=880 8.32 (5.94) 6.86 (3.76-11.45)	-1.19 (-1.76 to -0.64)*	<0.001
Mode of delivery				
Non operative vaginal delivery	1 150 (83.3)	1 140 (82.7)	1.01 (0.97 to 1.04) ‡	0.71
Caesarean delivery	143 (10.4)	148 (10.7)	0.96 (0.78 to 1.20) ‡	0.79
Operative vaginal delivery	88 (6.4)	91 (6.6)	0.97 (0.73 to 1.28) ‡	0.87
Emergency delivery	138/143 (96.5)	146/148 (98.6)	0.98 (0.94 to 1.01) ‡	0.42
Meconium stained amniotic fluid	233/1 238 (18.8)	320/1 127 (28.4)	0.66 (0.57 to 0.77) ‡	<0.001
Use of epidural anaesthesia	729 (52.8)	669 (48.5)	1.09 (1.01 to 1.17) ‡	0.03
Hypertensive disorders§	19 (1.4)	42 (3.0)	0.45 (0.26 to 0.77) ‡	0.004
Endometritis	18 (1.3)	6 (0.4)	3.00 (1.19 to 7.52) ‡	0.02

Values are numbers (percentages) unless stated otherwise
CI, confidence interval
*mean (95 % CI) difference between groups
†From start of active labour until delivery, time for cervical ripening excluded
‡Relative risk (95 % CI)
§Hypertensive disorders of pregnancy including eclampsia and haemolysis, elevated liver enzymes, low platelets count (HELLP)

Comments

Perinatal mortality and morbidity have been shown to increase already after 39 GW, but with a greater increase after 41 GW (Ingemarsson et al., 1997; Linder et al., 2017; Muglu et al., 2019; Nakling et al., 2006; Olesen et al., 2006; Reddy et al., 2006; Rosenstein et al., 2012). IOL has been proposed as a solution to decrease perinatal mortality and morbidity in women in late term and postterm

pregnancies (Hussain et al., 2011; Middleton et al., 2020; Wennerholm et al., 2009). Our study is in agreement with these findings regarding perinatal mortality. However, we could not show a difference in composite perinatal morbidity and mortality as was presented in the INDEX trial (n=1 801) where a significant decrease was reported from 3.1 % to 1.7 % (p-value 0.045) in the IOL versus expectant management group (Keulen et al., 2019). The INDEX trial did not show a significantly decreased perinatal mortality (one death in the IOL group and two in the expectant management group).

In our study routine fetal surveillance with CTG or ultrasound assessment of biophysical profile was not performed. Fetal surveillance was performed according to local protocol and differed between centres. In the Stockholm region an ultrasound was performed at 41 GW before recruitment to the study. It can be argued that the lack of *routine* fetal surveillance might result in fetuses at risk not being detected and therefore contribute to the high perinatal mortality in our study. However, the rate of adverse perinatal outcome was not higher in our expectant management group compared with the one in the INDEX trial (2.2 % and 3.1 %, respectively). In addition, the gestational age at delivery was higher (292 days) in our expectant management group than in the INDEX trial (289 days) resulting in a higher rate of women at risk for stillbirth in the expectant management group in our trial. This might contribute to the higher mortality rate in our study.

Furthermore, the decrease in perinatal mortality was achieved without an increase in caesarean and operative vaginal deliveries. Neither did the other maternal morbidities increase with IOL management.

The present study has a major strength due to its size and the fact that, even though only a small part of eligible women were recruited (22 %), the participating women represent a Swedish low-risk pregnant population. This was evident when the study population was compared with the Swedish background population. However, the early termination of the trial due to safety reasons is a limitation and results in an uncertainty of the magnitude of the reduction of perinatal mortality (Bassler et al., 2010). Yet, the early termination of the trial does not affect other outcomes, such as e.g. admission to neonatal care, meconium stained amniotic fluid, rate of small for gestational age newborns and hypertensive disorders of pregnancy that were all in favour of IOL.

Paper II

Induction of labour at 41 weeks or expectant management until 42 weeks: a systematic review and an individual participant data meta-analysis of randomised trials.

For this IPD-MA we performed a systematic literature search for RCTs comparing IOL at 41+0-2 GW with expectant management until 42+0-1 GW and found three trials. The three RCTs included 5 161women (n= 600, n=1 801 and n=2 760) (Gelisen et al., 2005; Keulen et al., 2019; Wennerholm et al., 2019). Two of the trials contributed with IPD (n=4 561) (Keulen et al., 2019; Wennerholm et al., 2019). The two trials included in the IPD-MA had mostly low-risk of bias (Table 8).

In total, 79.8 % (1 821/2 281) of the women in the IOL group were induced, whereas labour started spontaneously in 19.9 % (455/2 281) of the women. In the expectant management group 30.4 % (694/2 280) of the women were induced, while labour started spontaneously in 69.5 % (1 584/2 280). Scheduled caesarean delivery was performed in 0.2 % of the women (5/2 280) in the IOL group and 0.1 % (2/2 281) in the expectant management group.

Table 8. *Risk of bias within individual RCTs included in the individual participant data meta-analysis*

Author	Gelisen et al.	INDEX trial	SWEPIS trial
Selection bias	Unclear	Low	Low
Performance bias	High*	Moderate*	High*
Detection bias	Unclear	Low	Low
Attrition bias	Low	Low	Low
Reporting bias	Unclear	Low	Low
Conflict of interest bias	Low	Low	Low

*The lack of blinding in all RCTs are due to the nature of intervention i.e. it is not possible to blind the participants and staff

Perinatal outcome

The primary perinatal adverse composite indicator outcome was significantly lower in the IOL group as was perinatal mortality, compared with the expectant management group (Table 9). Number needed to treat (NNT) was 175 (95 % CI 94 to 1 267) and 326 (95 % CI 177 to 2 014), respectively. Furthermore, admission to neonatal care and neonatal care ≥4 days were significantly lower

in the IOL group with a NNT of 103 (95 % CI 59 to 385) for neonatal care ≥4 days. Newborns in the IOL group weighed significantly less than newborns in the expectant management group and were not as prone to have macrosomia. Other perinatal outcomes did not differ between groups.

The meta-analysis on aggregated data from all three trials regarding perinatal mortality did not differ much from the IPD-MA. Perinatal mortality was significantly lower in the IOL group compared with expectant management group (Figure 9).

Table 9. *Statistically significant perinatal outcomes in the population included in the individual participant data meta-analysis*

Variable	Induction group (n=2 281)	Expectant management group (n=2 280)	Relative risk/ Peto odds ratio (95 % CI)	p-value
Primary composite outcome*	10 (0.4)	23 (1.0)	0.43 (0.21 to 0.91) †	0.027
PNM‡	1 (0.0)	8 (0.4)	0.21 (0.06 to 0.78) #	0.019
Stillbirth	1 (0.0)	7 (0.3)	0.22 (0.06 to 0.89) #	0.034
Admission to a neonatal care§	79 (3.5)	109 (4.8)	0.72 (0.54 to 0.96) †	0.024
Admission to a neonatal care ≥4 days	24 (1.1)	46 (1.9)	0.52 (0.32 to 0.85) †	0.009
Birth weight (g) Mean (SD)	3 764 (417)	3 823 (439)		<0.001
Macrosomia (≥ 4500 g)	92 (3.9)	155 (6.7)	0.59 (0.46 to 0.76)†	<0.001

Values are numbers (percentages) unless stated otherwise
CI, confidence interval; IPD, individual participant data; MA, meta-analysis; PNM, perinatal mortality; SD, standard deviation. Relative risk is adjusted for RCT. P-value correspond to the method used to calculate the relative risk/odds ratio
*Including perinatal mortality, Apgar<4 at five minutes, HIE II-III, intracranial haemorrhage, neonatal convulsions, meconium aspiration syndrome, obstetric brachial plexus injury, mechanical ventilation within 72 hours
†Adjusted relative risk
‡Stillbirth and neonatal mortality (live births with mortality <28 days)
#Peto odds ratio
§Neonates admitted only for routine observation excluded

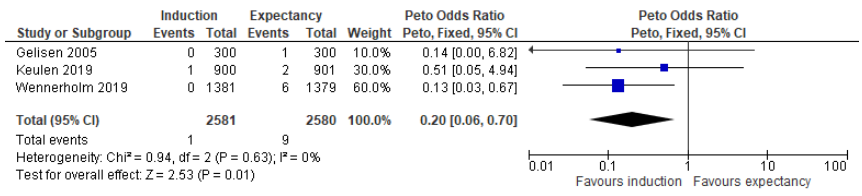

Figure 9. *Meta-analysis on aggregate data regarding perinatal mortality*

Maternal outcome

Hypertensive disorders of pregnancy occurred significantly less frequently in the IOL group compared with the expectant management group with a NNT of 57 (95 % CI 39 to 106) (Table 10). The only other maternal outcome that differed was use of pain treatment (epidural/spinal anaesthesia or opiates) and it was significantly higher in the IOL group (Table 10).

Table 10. *Mode of delivery and statistically significant maternal and delivery outcomes in the population included in the individual participant data meta-analysis*

Variable	Induction group (n=2 281)	Expectant management group (n=2 280)	Relative risk (95 % CI)	p-value
Pain treatment (Use of epidural/spinal/opiates)*	1 153 (50.5)	1,058 (46.4)	1.09 (1.03 to 1.16)	0.005
Use of epidural anaesthesia	998 (43.8)	906 (39.7)	1.10 (1.03 to 1.17)	0.006
Use of opiates	184 (8.1)	173 (7.6)	NE	NE
Meconium stained amniotic fluid	380/2 138 (17.8)	525/2 028 (25.9)	0.68 (0.61 to 0.77)	<0.001
Use of oxytocin†	1 440 (63.1)	1 077/2 280 (47.2)	1.33 (1.26 to 1.40)	<0.001
Mode of delivery	n=2,281	n=2,280		
Spontaneous vaginal delivery	1 860 (81.5)	1 836 (80.5)	1.01 (0.98 to 1.04)	0.41
Caesarean delivery	240 (10.5)	245 (10.7)	0.98 (0.83 to 1.16)	0.81
Operative vaginal delivery	181 (7.9)	199 (8.7)	0.91 (0.75 to 1.10)	0.33
Hypertensive disorders‡	26 (1.1)	66 (2.9)	0.39 (0.25 to 0.61)	<0.001

Values are numbers (percentages) unless stated otherwise. CI, confidence interval; HELLP, haemolysis, elevated liver enzymes, and low platelet count; IPD, individual participant data; MA, meta-analysis; NE, not estimated due to zero events in both arms in SWEPIS
*In the INDEX trial a combination of epidural and opiates was possible
†Both induction and/or labour augmentation
‡ Hypertensive disorders of pregnancy including eclampsia and haemolysis, elevated liver enzymes, low platelets count (HELLP)

Delivery outcomes

Meconium stained amniotic fluid was significantly less frequently present in the IOL group and use of oxytocin was significantly more frequent compared with the expectant management group (Table 10). Mode of delivery was similar in both groups and the main indication, in both groups, for caesarean delivery was failure to progress (50.0 % [120/240] in the IOL group and 49.8 % [122/245] in the expectant management group).

Subgroup analysis

We found an interaction with a p-value of 0.01 in the treatment effect for parity and the primary adverse perinatal composite indicator outcome. The risk of an adverse perinatal composite indicator outcome was significantly reduced in the IOL group compared with the expectant management group in nulliparous but not in parous women. NNT in nulliparous women were 79 (95 % CI 49 to 201). There were no other significant interaction effects for BMI, maternal age or fetal sex with the primary outcome or caesarean delivery.

Comments

Our IPD-MA showed a significant decrease in the primary adverse perinatal composite indicator outcome as well as perinatal mortality without increasing maternal morbidity such as e.g., caesarean delivery, postpartum haemorrhage and perineal lacerations III and IV. These findings agree with the Cochrane review (Middleton et al., 2020) and several other cohort studies (Grunewald et al., 2011; Lidegaard Ø 2020; Lindegren et al., 2017; Stock et al., 2012; Zizzo et al., 2017). However, there are other large observational studies that report an increase in caesarean deliveries, but a decrease in neonatal morbidity (Pyykonen et al., 2018) or no differences in perinatal outcome or caesarean deliveries (Rydahl, Declercq, et al., 2019). Furthermore, we showed in a subgroup analysis that newborns of low-risk nulliparous women will benefit substantially from IOL at 41 GW, whereas no such effect was found in multiparous women. This agree with some studies (Ingemarsson et al., 1997; Lindegren et al., 2017), but not all (Lawn et al., 2016; Middleton et al., 2020).

A strength of this study is that we performed an IPD-MA which is considered to have the highest certainty of evidence value of all different study designs (Stewart et al., 2002). However, a limitation is the low number of included trials and women. This increases the risk that one trial's limitations will have a substantial impact on the result of the IPD-MA. Furthermore, the low number of included women limits the possibility of exploring rare severe perinatal and maternal outcomes separately.

Paper III

Induction of labour at 41 weeks of gestation versus expectant management and induction of labour at 42 weeks of gestation: A cost-effectiveness analysis.

This CEA alongside SWEPIS included 2 746 of the 2 760 included women in SWEPIS. Hence, costs for eight women in the IOL group and six in the expectant management group could not be retrieved. All of these women had an average length of hospital stay and the rate of caesarean deliveries were similar I both groups. None of the newborns were admitted to neonatal care.

Total costs

The total cost per birth between the groups were similar, but all components of the total cost did significantly differ individually. In the robustness analysis, however, the total birth cost between the groups differed significantly and were more expensive in the IOL group (Table 1 in Paper III).

Incremental cost-effectiveness ratio

The ICER per LY and QALY with IOL at 41 GW compared with expectant management was slightly higher with a 95 % CI ranging from being dominant, i.e., lower costs and better health outcomes, to slightly higher costs. The results from the robustness analysis, based on cost data from the Sahlgrenska University Hospital, showed a 4.5 times higher cost with IOL indicating a higher ICER than the main analysis (Table 2 in Paper III).

In a cost-effectiveness plane 74 % of the ICERs ended up in the upper right quadrant indicating that IOL gave a better health outcome, but were more expensive, and 25 % of the ICERs ended up in the lower right quadrant indicating better health outcome and less costly (Figure 2 in Paper III).

Comments

Our study presented higher LYs and QALYs in the IOL group based on the reduction in perinatal mortality. The ICER had a cost level below €1 000 and a 95 % CI ranging from less expensive and better (dominant) to costs under €5 000. These costs are well below what both the Swedish National Board of Health and Welfare (€50 000) and National Institute for Health and Care Excellence in the UK (€21 750 – 32 635) considers to be high costs per QALY (NICE, 2013; SBU, 2018). Within the robustness analysis ICERs for LYs and QALYs were also below these thresholds by a large margin. Furthermore, since our ICER is dependent on the reduction in perinatal mortality, which might be overestimated in SWEPIS, we did a calculation where only one perinatal death differed between groups. The ICER was still far below the threshold for what considers to be cost-effective according to both the Swedish National Board of Health and Welfare and National Institute for Health and Care Excellence in the UK.

Our findings are supported by studies published earlier. The only two cost-effectiveness studies exploring IOL at 41 weeks compared with expectant management both show IOL to be cost-effective (Goeree et al., 1995; Kaimal et al., 2011). In addition, a health economic analysis alongside of the ARRIVE trial (Grobman et al., 2018), including 20 % of the study population, concludes that total costs are similar in both groups indicating a policy of IOL, compared to expectant management, will not increase costs (Einerson et al., 2020).

Being a cost-effectiveness study alongside a large RCT conducted in the real delivery/neonatal care context with a low number of lost to follow up was a major strength of this study. A limitation was the short follow up time and lack of information on additional health-costs after discharge from hospital.

Paper IV

This study evaluated 1 213 women in need of cervical ripening out of 1 638 induced women within SWEPIS (Wennerholm et al., 2009). Out of the 1 213 women, 744 were induced with OM and 469 with TVBC. The choice of method was decided by the obstetrician in charge and women with a low

Bishop score were more prone to receive OM while women with a high Bishop score were more frequently induced with TBVC.

Efficacy outcomes

Our primary outcome regarding efficacy 'vaginal delivery within 24 hours' was significantly lower in the OM group compared with the TBVC group. The same applied to vaginal delivery within 36 hours. Time-to-vaginal delivery and time-to-delivery were both significantly longer in the OM group compared with the TBVC group (Table 2 and 3 in Paper IV).

Oxytocin (both induction and/or labor augmentation) was significantly less used in the OM group compared with the TVBC group. Non-operative vaginal delivery, caesarean delivery, operative vaginal delivery or using more than one cervical ripening agent did not significantly differ between groups. Indications for caesarean delivery were comparable in both groups (Table 2 in Paper IV).

Safety outcomes

Neither the primary neonatal or maternal safety composite outcome differed significantly between groups. The secondary neonatal outcome metabolic acidosis was significantly higher in the OM group versus the TVBC group. Fever and therapeutic intravenous treatment with antibiotics during labor were significantly less frequent in the OM group compared with the TVBC group. However, the number of women with sepsis or endometritis were similar in both groups. None of the other secondary neonatal or maternal outcomes differed significantly between the groups (Table 4 in Paper IV).

Women's childbirth experience

We had an overall high response rate for both CEQ 2.0 (72.0 %) and VAS (78.3 %). The CEQ 2.0 total score was comparable in both groups. The childbirth experience measured with VAS did not differ between groups, nor

did the rate of women estimating the childbirth experience as very good (VAS 8-10) or as very bad (VAS 1-2) (Table 5 in Paper IV).

Comments

These efficacy findings are comparable with another large Swedish cohort study (Wollmann et al., 2017) comparing TVBC with both OM and vaginal dinoprostone. Furthermore, a Swedish RCT in full term pregnancies comparing TVBC, vaginal misoprostol and dinoprostone also report TVBC to be more effective (Prager et al., 2008). In addition, a Turkish RCT on late term and postterm pregnancies comparing TVBC and vaginal misoprostol reports in favour of TVBC regarding efficacy (Gelisen et al., 2005). However, the latest Cochrane review does not support our findings (de Vaan et al., 2019), but only two trials (n=602 [India]) and n=180 [Sri Lanka]) formed the basis of the conclusion in favour of OM regarding efficacy (Mundle et al., 2017; Somirathne et al., 2017). The largest one included a high-risk population (women with hypertensive disorder, IOL starting at 28 GW) and the smaller one included a low-risk population (women at 40+6 GW). These conflicting results can possibly be explained by different populations, different indications for IOL, different dosages and intervals of misoprostol, different volumes in the balloon or maximum duration of balloon treatment.

Our study lacked power to investigate safety outcomes. In contrast to two RCTs (Mundle et al., 2017; Ten Eikelder et al., 2016) we found significantly more metabolic acidosis in the OM group. Furthermore, also in contrast to earlier published studies, we report less fever and less use of antibiotics during labour in the OM group (Aghideh et al., 2014; Ten Eikelder et al., 2016).

In contrast to some qualitative and cohort studies (Falk et al., 2019; Hildingsson et al., 2011; Lou et al., 2019), our study showed an overall positive childbirth experience in both groups. In addition, Mundle et al., a RCT from India comparing IOL with OM and TVBC in women with hypertensive disorders reported more women being satisfied with their IOL in the OM group (Mundle et al., 2017). However, these results are difficult to compare with ours due to different indication for IOL and different cultural context.

A strength of this study was the included population, consisting of low-risk women exclusively, in combination with the comparison between OM and TVBC. There are few studies evaluating OM compared with TVBC in late term and postterm pregnancies (Goonewardene et al., 2014; Somirathne et al., 2017). In addition, a strength was that we report on women's childbirth experience. The study was limited by the study design. A cohort study always includes residual confounders. In this study the induction methods were used at markedly different rates and applied somewhat differently in the participating centres, which was a limitation. Hence, the effect of centre was difficult to fully adjust for in our analysis.

GENERAL DISCUSSION

The aim of this thesis was to clarify if a policy of IOL at 41+0-2 GW was superior, in terms of neonatal and maternal outcomes and health economic aspects, as compared with expectant management and induction at 42+0 GW in healthy women with a low-risk singleton pregnancy as well as exploring different cervical ripening methods. Our results show a reduction in adverse perinatal outcomes without an increase in maternal adverse outcomes or caesarean deliveries with a policy of IOL at 41+0-2 GW compared with expectant management and IOL at 42+0-1 GW. Hence, our results are reassuring for those who already changed management from IOL at 42+0 GW or beyond to IOL at 41+0 GW.

Perinatal and maternal outcomes

In Paper I we showed a decrease of perinatal mortality in the IOL group, but not of the adverse perinatal composite indicator outcome or the composite outcome including only the morbidity components. However, in Paper II we showed both a decrease in perinatal mortality alone and in the adverse perinatal composite indicator outcome. Furthermore, in Paper II a subgroup analysis indicated that nulliparous women will benefit from IOL at 41 weeks. For parous women we could not demonstrate an effect. However, the incidence of adverse perinatal outcomes was very low in parous women and if there would be a benefit/harm for them the number needed to treat/harm would be substantially higher in order to prevent/cause one event of adverse outcome. The overall reduction in adverse perinatal outcome was achieved without increasing caesarean or operative vaginal deliveries or increasing maternal morbidity.

The decrease in perinatal mortality without increase in caesarean deliveries are in line with most MA published (Alkmark, Berglin, et al., 2020; Hussain et al., 2011; Middleton et al., 2020; Wennerholm et al., 2009) but not all (Rydahl, Eriksen, et al., 2019). The Rydahl et al. publication is a MA consisting of a combination of RCTs, quasi randomised studies (retrospective observational studies) and observational studies and the authors conclude they do not have the power to evaluate perinatal mortality. In contrast to our findings the most

recent Cochrane review, including 34 RCTs, could not find any differences in perinatal mortality or stillbirth for timing of IOL according to parity (Middleton et al., 2020). However, observational studies have shown an increased risk of PNM in nulliparous women with increased gestational age, but not in parous women (Ingemarsson et al., 1997; Lindegren et al., 2017).

Regarding caesarean delivery, in the Rydahl MA, the pooled estimate from RCTs of caesarean delivery is a mixture of RCTs and quasi randomised studies (Rydahl, Eriksen, et al., 2019). The two true RCTs included in the MA reported similar rates of caesarean deliveries in both groups (Gelisen et al., 2005; Heimstad, Skogvoll, et al., 2007) Caesarean did not differ according to parity in the latest Cochrane review, nor did admission to NICU, operative vaginal delivery or lacerations III and IV (Middleton et al., 2020).

We present other important differences between the IOL and the expectant management groups in both Paper I and II, such as a decrease in admission to neonatal care and macrosomia in the IOL group compared with the expectant management group. In Paper II we investigated the outcome 'admission to neonatal care ≥4 days' in order to mirror only severe neonatal outcomes, and we showed a significant decrease in the IOL group. In addition, in both Paper I and II, we present a significant decrease in hypertensive disorder of pregnancy in the IOL group compared with the expectant management group.

Cost-effectiveness outcomes

In Paper III a CEA is presented and we found no significant difference of total cost per birth for a policy of IOL at 41+0 GW compared with expectant management. However, the delivery cost was significantly higher in the IOL group. The difference was €342 (95 % CI 184-501), a 13 % increase. The costs for neonatal care and costs for antenatal outpatient's visits were significantly lower in the IOL group. The LYs and QALYs in the IOL group were higher, due to the reduction in perinatal mortality. The ICER for a policy of IOL at 41+0 GW was low (€601) in relation to what the Swedish National Board of Health and Welfare considers a threshold for high costs per QALY (€50 000) (SBU, 2018). These findings concur with earlier studies (Einerson et al., 2020; Goeree et al., 1995; Kaimal et al., 2011). In Walker et al. they even report a

lower cost per birth in the IOL group and an ICER of –£114 500 per QALY (Walker et al., 2017).

According to our findings, implementing IOL at 41+0 GW would involve IOL of approximately 22 % of the pregnant population in Sweden. An increase in delivery cost per women by €342 would result in an additional cost of €8 652 600 (115 000 deliveries per year in Sweden) for the delivery care system. However, this increase in costs will be compensated by a decrease in costs for outpatient visits and costs for neonatal care.

In summary, the intervention IOL at 41+0-2 GW is cost-effective by far according to the Swedish National Board of Health and Welfare and the possible increase of costs for the delivery units will be compensated by a decrease in costs for the antenatal clinics and neonatal units.

Methods for cervical ripening

The need for the safest and most effective method for ripening the cervix in this low-risk late term population will be of even more interest when as much as 22 % of the pregnant population could undergo IOL. In Paper IV we evaluated OM compared with TVBC as a method for ripening the cervix. The primary efficacy outcome was significantly lower in the OM group. Hence, more women in the TVBC group gave birth within 24 hours and these results persisted at 36 hours, but not at 48 hours. In the literature, we only found two RCTs comparing OM with TVBC in low-risk almost (40+6 GW) late term pregnancies. These relatively small trials (n=152 and n=180) were from the same research group (Goonewardene et al., 2014; Somirathne et al., 2017). In the trial from 2014, a regime of two doses of 25 µg of OM four hours apart was compared with a regime of TVBC for 24 hours. There were no differences in induction-to-delivery interval or caesarean delivery rate between the two groups, but significantly more women (both nulliparous and parous) had a Bishop sore > 6 in the TVBC group after 24 hours (Goonewardene et al., 2014). The trial from 2017 compared three doses of 50 µg of OM four hours apart with 24 hours' treatment with a TVBC. Significantly more women (both nulliparous and parous) gave birth vaginally within 24 hours in the OM group compared with the TVBC group. There were no differences in caesarean

delivery rate between the groups (Somirathne et al., 2017). The latest Cochrane review evaluating TVBC compared with OM included eight RCTs (including the two trials described above), but only two (n= 602 and 180) reported on the primary outcome 'vaginal delivery not achieved within 24 hours' (de Vaan et al., 2019). The primary outcome was in favour of OM. The largest trial (n= 1 845) included in the review, but not in the primary outcome, also presented results in favour for OM regarding efficacy (Ten Eikelder et al., 2016). They reported delivery within 24 hours and not vaginal delivery within 24 hours, hence this was probably the reason why it was not included in the analysis of the primary outcome. Our findings are in line with a large cohort study conducted in a Swedish population including both high and low-risk term and postterm singleton pregnancies (Wollmann et al., 2017). Furthermore, our study shows similar results as two RCTs comparing vaginal misoprostol with TVBC (Gelisen et al., 2005; Prager et al., 2008).

Our study was not powered for evaluating safety outcomes and we only found a significant difference in metabolic acidosis in favour of TVBC. However, there were several missing results of the umbilical cord blood sampling reducing the reliability of the results. Ten Eikelder et al. presented similar safety results in the OM and the TVBC group (Ten Eikelder et al., 2016). The Cochrane review concluded that it is still unclear whether there is any difference between the two methods regarding safety for the newborns and the women (de Vaan et al., 2019).

In Paper IV we also investigated women's experience regarding the two different cervical ripening methods. Overall women reported a positive childbirth experience and the results were similar in both groups.

Methodological considerations

The basis of this thesis was a large RCT (SWEPIS) followed up by an IPD-MA and a health economic evaluation using a CEA method. The two principal cervical ripening methods used within SWEPIS, OM and TVBC, were assessed by a cohort study. However, the best study design for evaluation of rare outcomes, such as perinatal mortality and morbidity have been up for debate.

Randomised controlled trials

In order to establish a causal relationship between an intervention and outcome RCTs have been considered the golden standard methodology because their ability to minimise risk of bias regarding known and unknown confounders. However, conducting RCTs, especially when a rare outcome is studied is costly, time-consuming and the external validity may be questionable. In addition, when studying a rare outcome, the large number of patients needed to be recruited will also be an obstacle hard to overcome. Hence, observational studies based on large databases "big data" (Greenwood, 2014) have been promoted to replace RCTs in these circumstances (Franklin et al., 2017; Galson, 2016). Franklin et al. and Hernán et al. both describe how to plan and conduct such observational studies in order to emulate the target trial as appropriately as possible (Franklin et al., 2017; Hernán et al., 2016). Moreover, a controversy regarding observational studies ability to estimate effect size is ongoing. There is evidence for observational studies, especially retrospective observational studies, overestimating the effect size (Ioannidis et al., 2001; Oliver et al., 2010). Yet, it has been argued that this is not true and that well executed observational studies does not differ regarding effect size compared with RCTs (Concato et al., 2000; Golder et al., 2011). However, there are still difficulties to overcome and large observational studies may incorrectly conclude that there is no effect of the intervention or inaccurately promote a causal effect where there is none (Collins et al., 2020). Furthermore, even though adjustments for known covariates are made the treatment effect may still be attributed to both unmeasured and unknown confounders.

A difficulty often discussed when assessing RCTs is the external validity. External validity refers to how applicable the results of a study is outside the study's context. That is, are the results of the current trial generalizable to a broader population? Roberts et al. argue that it is more important to focus on the mechanism of the intervention than if the population studied is statistically representative or not (Roberts et al., 2014). Hence, one should not confuse statistical inference with scientific inference. In statistical inference, a representative sample is important for the conclusion to be more accurate. Regarding scientific inference, understanding how biology functions in interaction with the intervention is more important (Roberts et al., 2014)

(Rothman et al., 2013). Therefore, when evaluating external validity consideration should be taken to the mechanism leading to the treatment effect in relation to the study populations characteristics and ask if it is reasonable that the treatment will have a similar effect in another population and under other circumstances.

In summary, conducting a RCT generate a higher certainty of evidence than observational studies, even though the latter are based on larger cohorts.

Meta-analysis

As stated in the previous section of this thesis, conducting a large RCT in order to evaluate rare outcomes can be difficult. A method to overcome these problems is to conduct a MA where you combine evidence from several studies evaluating the same research question. Hence, a MA aims at summarising estimates of individual trials into a pooled estimate for all included trials. In order to perform a MA a systematic review of the literature is needed to gather all evidence within the research field.

MA can be conducted with aggregate data or with IPD (Riley et al., 2010). Aggregate MA is the most frequently used method due to its low-cost and time effectiveness as the data is easily accessible (data from published articles). In conducting an aggregate MA, a conversion of individual study results into a common scale is made (events per population). Furthermore, each individual study estimate is reported in a forest plot and finally a pooled estimate for all included studies is calculated. The forest plot enables a direct interpretation of estimates of the included trials and their contribution to the pooled estimate. However, MA of aggregated data is vulnerable to publication bias (research with negative results are less likely to be published) leading to a risk of overestimating effects (Blettner et al., 1999). The researcher conducting the MA has restricted control over the data, hence the heterogeneity between the included studies is of most importance when interpreting the results of an aggregate MA. The heterogeneity reflects the variation in study design, analysis approaches, how the outcome variables have been categorised and how the data have been collected (Blettner et al., 1999).

An advantage of an IPD MA includes a greater control of the data and the possibility to decrease the heterogeneity through harmonising definitions of outcomes, inclusion and exclusion criteria and measures used in the different studies (Stewart et al., 2015). An additional advantage is the possibility to include unpublished data by approaching all researchers within the field in order to increase the number of included participants and, hence, increase power of detecting differences in rare outcomes (Riley et al., 2010; Tierney et al., 2015). Furthermore, in addition to improving quantity and quality of data, an IPD MA approach increases the possibility to undertake subgroup analyses (Tierney et al., 2015). However, an IPD MA is time consuming and expensive due to its need for collaboration between all researchers responsible for the included trials, reanalyses and storage of data.

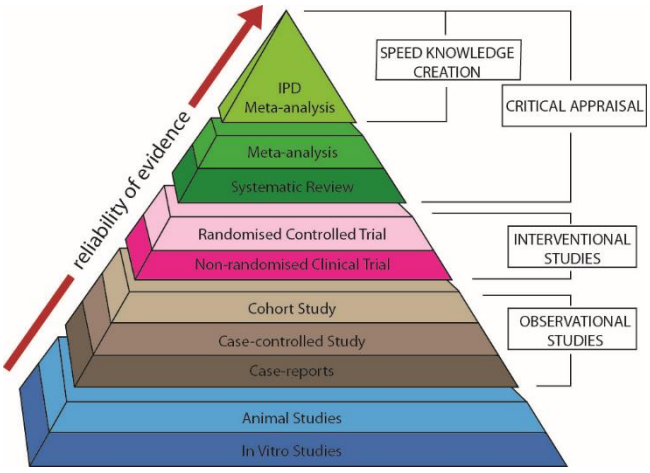

Figure 10. *Evidence pyramid. The hierarchy of evidence and the role of the individual participant data (IPD) meta-analysis in knowledge creation is presented. The reliability of evidence needed for the evidence informed decision making in health increases as we move up the pyramid. It is expected that individual participant data meta-analysis would speed up the knowledge creation. (Krleža-Jerić et al., 2016)*

In summary, a well-executed IPD-MA is considered to create more robust and detailed results, deliver the highest certainty of evidence and give more possibilities to investigate subgroups (Figure 10). However, it is more

expensive and time consuming to perform (Krleža-Jerić et al., 2016; Tierney et al., 2015; Tierney et al., 2020).

Individual participant data meta-analysis methods

When conducting an IPD-MA, consideration regarding the recruitment of individuals according to different trial protocols must be taken. To merge the IPD in to a mega file for analysis could lead to biased comparisons and overestimation of effects. Hence, you need to account for participants' origin of trial during the analysis. The two-step approach, in relation to the one-step approach, is more commonly used. In a two-step approach, the first step of analysis includes harmonising the data from each included trial and calculate new estimates for each outcome in each individual trial. In the second step these estimates are combined using the same methods as in an aggregate MA resulting in a forest plot. In the one-step approach a single model is adapted to the individual study results and typically includes a regression model where trial origin and/or important differences between studies are accounted for. The one-step model usually reduces the risk of bias if the trials included are small, the effect is large or outcomes studied are rare. Hence, a well-executed one-step approach is to prefer over the two-step approach because it does not need additional steps in order to make the most of the available data (Fanshawe et al., 2019; Tierney et al., 2015).

In order to investigate rare outcomes and perform subgroup analysis we preferred an IPD MA approach instead of a regular aggregate MA. In addition, a one-step approach for our IPD-MA was chosen because of the rarity of studied outcomes and the relatively large effect. However, for outcomes where Peto odds ratio was used a two-step approach was chosen because Peto odds ratio can only be used to pool odds ratio.

Health economic analyses

The method of choice regarding economic evaluation of interventions within the healthcare system is CEA where you measure outcome in e.g., LYs or QALYs gained or disability days prevented and total costs for an intervention (Sanders et al., 2016). Other approaches are to perform a cost consequence analysis, cost minimisation analysis or a cost benefit analysis (Husereau et al.,

2013). The cost consequence analysis investigates costs and consequences in a broader perspective without using LYs and QALYs. Costs and outcomes of a new intervention is presented in a descriptive table leaving the reader to decide on relevance and importance of the information (Mauskopf et al., 1998). A cost minimisation analysis is conducted when there are no differences in treatment effects between the interventions compared. Thus, only costs associated with respective intervention are compared (R. Robinson, 1993). Lastly, cost benefit analysis is a method where all consequences, both health and costs, are presented in monetary units. Hence, gain in health and lives are converted into a monetary unit based either on what a person is paid (the human capital approach) or on individuals' observed or stated favourable treatment (the estimate of money a person is willing to accept for an increased risk or is willing to pay for a service) (R. Robinson, 1993).

All these approaches can be conducted alongside a clinical trial or be based on a decision analytic model or both. In order to perform a valid health economic analysis alongside a clinical trial it should be a pragmatic trial. That is, the trial should be conducted in a real clinical situation and not in a special research facility, the control treatment should be the most relevant cost-effective treatment used, the trial should be sufficiently powered and the follow-up period should be long enough to include relevant costs (Svensson, 2019). The medical results of the trial and costs collected from administrative systems are the raw data for performing a health economic analysis, usually a CEA if LYs or QALYs have been evaluated within the trial (Svensson, 2019). If some or all of these criteria are absent, a decision analytic model can be used. The method is based on data from previously published studies, registers, administrative systems or even expert estimates in lack of better options. A health economic evaluation alongside a clinical trial meeting all the above mentioned criteria renders the most reliable results. However, a decision analytic model is of most help when no clinical trial is present or when the follow-up is too short to include relevant costs.

Observational studies

Paper IV is an observational study within SWEPIS. As a result of the investigated research questions nature, the large number of included centres

and the IOL methods used being according to local protocol several confounders were present. A confounder is a variable different from the intervention itself, which affects both the exposure and the outcome (Figure 11). Therefore, a confounder could be the real cause of an established relationship and not the intervention studied. For example, Bishop score may influence both the choice of cervical ripening method and the delivery outcome and if not adjusted for could be the real cause of TVBC being more effective than OM.

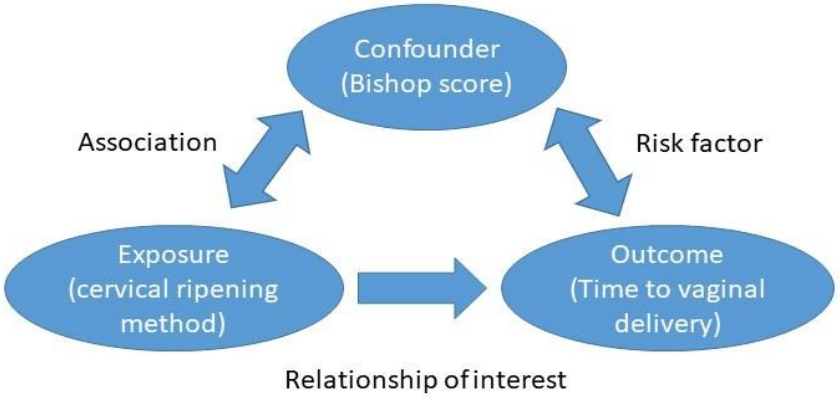

Figure 11. *Illustration of a confounder and its relationship with exposure and outcome*

A pragmatic criterion for conventional covariate adjustment in a regression model is to have at least 10 events per covariate included. This might be difficult when there are low numbers of events or few patients in an observational study (Harrell et al., 1984). Considering the large number of confounders in our study, conventional adjustment for covariates was not possible. Adjustment with PS is a way to overcome this obstacle. In a RCT the probability of allocation to either intervention is 0.5 but in an observational

study the probability is unknown and dependent on the reported baseline characteristics. A PS is the probability of intervention allocation based on the reported baseline characteristics including important confounders such as clinical environment and caretakers' preferences (Haukoos et al., 2015; Kuss et al., 2016). A regression model with the intervention as dependent variable and baseline characteristics before the intervention and/or known confounders as independent variables are often used to calculate the PS. Moreover, PS strives to convert an observational study to mimic a RCT as much as possible (Austin, 2011). Hence, we used the PS approach and included all centres (n=15) and most baseline characteristics in the process of creating the PS.

There are four different PS models to choose from, matched PS, stratified PS, inverse probability weighting PS and using the PS as a covariate in a regression model (Austin, 2011). What model to use depends on the distribution of the PS of the treatment and control group (Elze et al., 2017). In our study PS were distributed with poor overlap between the OM and TVBC group and a large number of PS was close to the value zero and one (Figure 12).

This distribution made it more difficult to choose the best PS model for our study. PS by matching would result in exclusion of a large part of the participating women because many women in the OM group would not get a match and therefore be excluded. PS by stratification would be a problem because possible strata would be too few due to the distribution but also due to low numbers of safety events in each stratum. PS by inverse probability weighting would give a substantial weight to a few women in both groups but especially in the OM group leading to more imprecise estimates of treatment effect and lack of covariate balance (Elze et al., 2017).

In Paper IV we choose to use the PS approach in order to construct intervention groups similar to each other enabling a more accurate evaluation of the relationship between the two interventions.

PS as a covariate in a regression model was the chosen method due to its simplicity and the pitfalls in the other PS models.

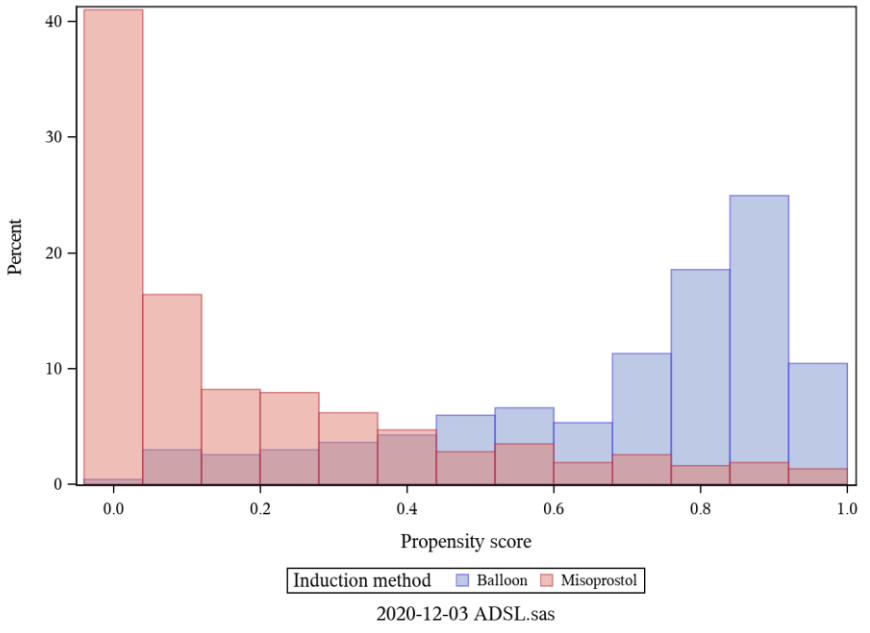

Figure 12. *Propensity score distribution in Paper IV*

Strengths and limitations

Strengths and limitations of the individual studies are presented in the result section under the subheading comments. Being able to perform a pragmatic RCT to evaluate rare outcomes is a strength. It was possible due to the collaboration of several delivery units in Sweden and the possibility to use national quality registers for collecting outcome data. Furthermore, an asset in this thesis is the low number of women and newborns lost to follow up increasing the reliability of the results. Lastly, evaluation of women's childbirth experience has been performed within SWEPIS (Wessberg et al., 2017, 2020) (Nilvér et al., 2021, Women's childbirth experience in the Swedish Postterm Induction Study [SWEPIS]: a multicentre, randomized, controlled trial, revision submitted Dec 2020), making it possible for us to investigate the differences in childbirth experience in Paper IV.

As stated in the result section regarding Paper I, a limitation affecting the results of Paper I-III is the early termination of SWEPIS. A study compared truncated with non-truncated RCTs and concluded that there was an increased risk of overestimating positive treatment effect/harm if an RCT is stopped early (Bassler et al., 2010). This risk increases if the outcome is rare, the number of patients included in the trial is low and if the effect size is small (Bassler et al., 2013). However, according to another research group this conclusion is invalid due to the fact that truncated studies are stopped early because a positive treatment effect/harm was achieved earlier than expected and therefore one can expect an average larger effect size than non-truncated trials (Goodman et al., 2010). On the other hand, simulation studies indicate that the overestimation effect is limited and exaggerated (Wang et al., 2016). Furthermore, the bias related to truncating a trial usually disappears when included in a MA (Bassler et al., 2010; Goodman et al., 2010). In Paper I, the outcome perinatal mortality is rare, which may increase the risk of overestimation. On the other hand, the number of included women are relatively high and the effect size is relatively large, which may decrease the risk of overestimation. Since SWEPIS has great weight in the IPD MA in Paper II, the risk of overestimation might not be completely neutralized. However, once again, the total number of patients is relatively large and the effect size is large regarding perinatal mortality, reducing the risk that the IPD MA overestimated the effect size due to that SWEPIS was stopped early. Furthermore, the results of SWEPIS and our IPD MA are similar to the results of the latest Cochrane review (Middleton et al., 2020). Thus, one can argue that the higher perinatal mortality in the expectant management group in Paper I is plausible as we know that the risk of fetal death increases with the duration of pregnancy and that the relative risk of fetal death is higher in late term/post term pregnancy (Muglu et al., 2019).

Another limitation of this thesis is the exclusion of women with a previous caesarean delivery or major uterine surgery. The caesarean delivery rate has increased considerably from 7 % in 1990 to 19 % in 2014 worldwide including both high-, middle- and low-income countries (Betrán et al., 2016). In Sweden, the increase is similar, in 1973 the caesarean rate was 5.3 % and in 2019 it had increased to 17.7 % (MBR, 2020). Between 1992 and 2015 a twofold increase of caesarean delivery with a previous uterine scar as indication was reported in Sweden (da Silva Charvalho et al., 2019). This situation results in a considerable number of women with a previous uterine scar not delivered at

41 GW. The latest Cochrane review assessing IOL compared with repeat elective caesarean delivery in women with a previous major uterine scar did not find any eligible RCTs (Dodd et al., 2017). Another Cochrane review comparing different IOL methods in women with a previous caesarean delivery found eight RCTs, all evaluating different methods (West et al., 2017). Presenting a pooled estimate of outcomes was not possible and most RCTs were underpowered for detecting differences in relevant outcomes.

A reason for not including women with a previous caesarean delivery in RCTs evaluating IOL compared to expectant management or repeat caesarean delivery is the relative contraindication for using prostaglandins as IOL method, due to an increased risk of uterine rupture. The risk of uterine rupture after IOL in women with a previous caesareans delivery is estimated to 1-2 % (ACOG, 2019). This causes methodological difficulties when planning a trial comparing IOL with expectant management or repeat caesarean delivery. Hence, including women with a previous uterine scar in a RCT in order to evaluate them as a subgroup would increase the need of included women significantly resulting in the RCT being impossible to perform. Including these women without the intention to evaluate them as a subgroup would make the interpretation of the results difficult. Therefore, we do not address the risks of IOL or expectant management for this group of women and their newborns in this thesis.

An additional limitation is the lack of patient involvement. It has been argued that patient and public involvement increases the relevance, validity, quality and success of research (Goodare et al., 1995; Gradinger et al., 2015) and several funders require patent and public involvement (Boivin et al., 2018). Furthermore, patient and public involvement in research has shown to increase patient enrolment in clinical trials (Crocker et al., 2018). Hence, when planning the next research project, we will consider patient and public involvement.

Ethical aspects

According to our research a policy change from IOL at 42+0 GW to 41+0 GW will reduce perinatal mortality and morbidity without increasing caesarean delivery, operative vaginal delivery or maternal morbidity such as postpartum

haemorrhage or perineal lacerations III-IV. Inducing women at 41+0 GW instead of 42+0 GW would potentially save 100 newborns per year in Sweden from severe morbidity or mortality. In addition, a reduction in admissions to neonatal care, newborns being small for gestational age or macrosomic and neonatal jaundice may be a result of early IOL. Furthermore, it is likely to decrease the number of women developing hypertensive disorder of pregnancy. Negative medical effects of a policy change may be an increase in use of pain relief and antibiotics during labour. However, overall, IOL in late term pregnancy is in line with the ethical principle of beneficence and non-maleficence.

The ethical principal of autonomy is important to consider. There are presently diverse and strong opinions among caregivers regarding IOL and its efficacy and safety. The absolute risks of mortality or severe morbidity are low with expectant management and IOL is associated with need for more pain relief and use of oxytocin (Alkmark, Keulen, et al., 2020; Keulen et al., 2019; Wennerholm et al., 2019). Furthermore, IOL is an intervention in the natural birth process and some women would like to avoid interventions and believe that the individual benefits being greater with expectant management. Hence, it is of utmost importance that written information on benefits and risks regarding IOL versus expectant management is unbiased, understandable and easily available for all women considering IOL at late term. In order to fulfil the ethical principal of autonomy it is crucial that women have the opportunity to make an informed decision.

The last of the four basic ethical principles within the healthcare system is justice. An increase in IOL can potentially create displacement effects resulting in women being referred to other hospitals or receiving substandard care. In addition, a policy of IOL at 41+0 GW would result in increased costs for the delivery unit, but not the total cost for the health care system. This might result in other health care services within the delivery care receiving reduced funding. However, the CEA in Paper III showed a cost per QALY far below the threshold for what is considered a cost-effective intervention. If there is a benefit in saved lives a higher cost would be justified.

In summary, there are no or little evidence supporting that offering IOL at GW 41+0 instead of expectant management until GW 42+0 should be unethical as long as the autonomy of the patient is respected.

CONCLUSION

This thesis has evaluated IOL at 41+0-2 GW as compared to expectant management and IOL at 42+0-1 GW with respect to perinatal, maternal and delivery outcomes. We have shown a benefit, especially in nulliparous women, for a policy of IOL at 41+0-2 GW due to:

- An overall reduction in perinatal adverse composite indicator outcome including perinatal death and severe neonatal morbidity. In a subgroup analysis these benefits were shown in nulliparous women. For parous women the effect is still unclear. The magnitude of risk reduction of perinatal mortality remains uncertain because the trial in Paper I was stopped early due to safety reasons.

- An overall reduction in admission to neonatal care, and especially a reduction of admission to neonatal care for four days or more indicating a reduction in newborns with severe morbidity.

- A reduction in women with hypertensive disorders of pregnancy.

- No increase in caesarean deliveries or operative vaginal deliveries.

- No increase in maternal morbidity such as postpartum haemorrhage or III- and IV-degree perineal lacerations.

- No significant increase in total costs for the health care system. ICER for both LYs and QALYs would be well below the recommended threshold for being cost-effective.

In Paper IV we showed an advantage of cervical ripening with a TVBC over the use of OM in late and postterm pregnancies due to:

- An increase in vaginal delivery within 24 hours when a TVBC is used for IOL. The results withstand at 36 hours.

- Shorter time from start of IOL to vaginal delivery in the TVBC group compared with OM.

- Similar results regarding caesarean and operative vaginal delivery.

- Similar results regarding neonatal and maternal safety outcomes.

- Similar results regarding women's experiences in the two groups.

In summary, there are medical benefits of IOL at 41+0-2 GW compared with expectant management and IOL at 42+0 GW. The health economic evaluation shows that IOL at 41 GW is cost-effective, i.e. entails a low cost per QALY gained. Furthermore, OM and TVBC are almost completely comparable, except for TVBC being slightly more effective regarding efficacy. Hence, it is of importance that women approaching 41+0 GW receive unbiased and easily understandable information, both oral and written, regarding benefits and risks with IOL at 41+0 GW compared with expectant management. It is a prerequisite in order to make an informed decision for herself and her unborn infant.

FUTURE PERSPECTIVES

Our four papers have contributed further information to the research question about how to manage late term pregnancies. Nevertheless, there are still scientific knowledge gaps. The extent of decrease in perinatal mortality in late term pregnancies is still not established. However, the likelihood of a new RCT powered to detect differences in perinatal mortality between IOL at 41 GW and expectant management until 42 GW being conducted is low.

Women with a previous major uterine scar

One important knowledge gap is, as mentioned earlier, the lack of information regarding management of women with a major previous uterine scar. The latest Cochrane review comparing elective repeat caesarean delivery with IOL in women with previous caesarean delivery found no RCTs addressing this research question (Dodd et al., 2017). Hence, recommendations regarding management strategies are based on non-randomised trials and benefits/harm with either strategy should be interpreted with caution. It would be of utmost interest to conduct a large RCT evaluating the risks with IOL at 41+0 GW compared with expectant management or elective repeat caesarean delivery. Several national guidelines do not consider IOL after one caesarean delivery as a contraindication, but stresses the importance that information regarding risks/benefits with IOL and elective repeat caesarean delivery is given to the women (ACOG, 2019; G. NICE, 2008; Sentilhes et al., 2013). In addition, most clinical guidelines advice against use of misoprostol for IOL in women with previous uterine scar due to a possible high risk of uterine rupture. As mentioned earlier the latest Cochrane review investigating different IOL methods in women with a previous vaginal delivery concludes there are not enough evidence from RCTs to promote one optimal method of IOL, but points out that one study using misoprostol was stopped early due to safety reasons (West et al., 2017). The likelihood of a large well executed RCT powered for investigating perinatal outcome being conducted is low due to the difficulties associated with conducting RCTs described earlier in this thesis.

Exploring other risk factors

As presented earlier in this thesis, advanced maternal age, high BMI, being nulliparous, and carrying a male fetus has been put forward as risk factors for stillbirth. We explored maternal age ≥35 years, parity, BMI ≥30 and fetal sex

in a subgroup analyses and found that parity was significant for adverse perinatal composite indicator outcome. We showed that nulliparous women benefit from IOL at 41+0-2 GW but the risk for parous women to be affected by an adverse perinatal outcome was considerably lower. However, an interaction analysis for perinatal death was not possible due to the low perinatal mortality rate. In Paper I all the perinatal deaths occurred in nulliparous women but in the INDEX trial two out of three perinatal deaths occurred in parous women (Keulen et al., 2019; Wennerholm et al., 2009). Furthermore, the subgroup analysis in Middleton et al. 2020 did not find perinatal mortality to bee differently affected by IOL in parous vs. nulliparous women. Therefore, there is still some uncertainty about how to select women at risk in the low-risk population of parous women (Delaney et al., 2008; Lawn et al., 2016; Lawn et al., 2006).

Exploring the optimum time point for induction of labour in low-risk pregnancies

In planning and conducting our trial in Paper I we received questions regarding our choice of comparison groups. Opposing arguments for the optimal time point for delivery e.g. 38 GW, 39 GW or 40 GW have been put forward by several research groups (Grobman et al., 2018; Muglu et al., 2019; Rosenstein et al., 2012). Grobman et al. conducted a RCT comparing IOL at 39+0-4 GW with expectant management in over 6 000 nulliparous women and showed a nearly significant reduction in perinatal adverse outcome (a 95 % CI including 1.00 as the upper limit) and a significant reduction in caesarean deliveries (Grobman et al., 2018). Perinatal mortality and morbidity increases gradually after 38 GW but becomes progressively higher after 41 GW. A policy of IOL at term compared with late term would comprise approximately 50 % of all pregnancies. Hence, even though the rate of perinatal mortality and morbidity is lower at the 39-40 GW, the absolute number of fetuses and newborns with adverse perinatal outcomes are higher. Hence, conducting a trial comparing IOL at 39-40 GW with expectant management until 41 GW has the potential to save more fetuses and newborns from adverse outcomes. However, the NNT in order to avoid an adverse perinatal outcome would probably be high resulting in a need for a large study population in order to detect significant differences between the groups. Furthermore, it would increase the workload considerably at the delivery units.

Increased surveillance

Fetal surveillance with ultrasound or ultrasound in combination with CTG or CTG alone has been argued for in order to detect fetuses at risk. As described earlier in this thesis there is low evidence that any surveillance method reduces perinatal adverse outcome in late term pregnancy (ACOG, 2014; Delaney et al., 2008; Henrichs et al., 2019; Lindqvist et al., 2014; Nabhan et al., 2008). The possibility of performing a RCT comparing IOL at 41+0 GW with expectant management with different fetal surveillance methods and IOL at 42+0 GW seems unlikely with regards to our research results. However, the possibility to conduct a RCT comparing IOL in 39-40 GW with expectant management and different fetal surveillance methods or just comparing different fetal surveillance methods with no fetal surveillance might be feasible.

Organisational aspects

By tradition it is common, in Sweden today, to start IOL at the delivery ward and keep the woman there throughout the whole birth process until approximately two hours after delivery. Recent research shows a significant increase in time from admission to the delivery ward until delivery in induced women compared with those with spontaneous onset of labour (Grobman et al., 2018; Keulen et al., 2019; Wennerholm et al., 2019). If IOL at 41+0 GW is introduced in Sweden it will result in a considerable increase in women being induced and, thus, an increased burden on staff and room availability at the delivery units. It may compromise patient safety by displacement effects and an increase in referrals to other hospitals due to lack of available beds and staff. However, if current management strategy prevails some kind of fetal surveillance probably is needed (ultrasound assessment and or CTG) during the time period between 41 and 42 GW, which also requires increased resources.

In pursuance of maintaining safety, organisational changes have to be considered. The increase in inductions will consist of low-risk late term pregnancies with a fetus in cephalic position. Thus, one could argue that the need for fetal surveillance and nursing of the women might not be as important as when inducing a woman for other medical reasons. Hence, outpatient IOL could be an alternative. A recent Health Technology Assessment report regarding outpatient versus inpatient IOL in term low-risk singleton pregnancies concluded that there is very low certainty of evidence regarding

safety aspects for the mother and fetus (Berglin L, 2020). Further, they concluded there is a need for well conducted RCTs addressing this question. Another alternative could be to start the induction at a ward with less resources in order to spare the high-resource delivery ward and, thus, maintain patient safety.

Evaluation of policy change

If a policy change from IOL at 42+0 GW to IOL at 41+0 GW in women with a low-risk pregnancy is introduced in Sweden a golden opportunity to plan and perform a register-based cohort study will present itself. There have been policy changes regarding IOL in women with low-risk pregnancies at late term in other countries and several well executed cohort studies have derived from them. In Sweden, we have a long tradition of using data from national health data and quality registers for large cohort studies. Hence, a study comparing five years before the policy change, e.g., 2010-2015 with five upcoming years e.g., 2022-2027 would be of interest.

Summary

In summary, RCTs regarding how to manage women with a previous major uterine scar is warranted but probably difficult to conduct. Hence, we might have to rely on large cohort studies to gather more evidence. Moreover, continuous research on the optimal time point for IOL in low-risk pregnancies in order to reduce term perinatal mortality would be of interest, as would research on subgroups of low-risk women who benefit most from IOL. In addition, more research on effective fetal surveillance regimes that would benefit pregnant women and their unborn infant in term pregnancies would contribute to the understanding of the best management of low-risk pregnancies at term. Lastly, continuous evaluation of any policy change regarding IOL in low-risk late term pregnancies is of high interest.

ACKNOWLEDGEMENT

The completion of a thesis is not a one-man or woman job. Without help and encouragement from all the people involved I would not have reached the point I am at now. I would especially like to extent my gratitude to:

All women that participated in SWEPIS making this thesis possible.

Professor **Henrik Hagberg**, my main supervisor, for his patience, encouragement and endless support, making my research education possible. Thank you for giving me the opportunity to change research project and for sharing your vast knowledge in both experimental and clinical research.

Associate professor **Ulla-Britt Wennerholm**, one of my co-supervisors, for extensive knowledge, endless energy and always being available, even if it is past midnight. Thank you for giving me access to all your invaluable knowledge and for motivating me to finalise the thesis within the year 2021.

PhD **Ylva Carlsson**, my other co-supervisor, for constructive feed-back, patiently answering all my questions, and all you support when my motivation was running low.

Mattias Molin and **Per Ekman**, the Statistical Consulting Group, Gothenburg, for performing the statistical analyses.

Nils Crona, **Corinne Pedroletti**, **Anneli Falk** and **Henrik Almgren** who have all been Head of Clinical Department of Obstetrics at Sahlgrenska University Hospital during the years I have been conducting my research. Thank you all for believing in me and allowing me time off from clinical work.

All midwives and doctors in the SWEPIS study group, without your contribution my research would not have been possible.

The INDEX team for an educational and constructive collaboration. Thank you **Judit Keulen** and **Esteriek de Miranda** for challenging me and for the meaningful discussions we had.

All co-authors, thank you for creative discussions and sharing your knowledge to reach the published results.

Klaus Groth, Randa Akouri and **Helena Kopp Kallner** for a valuable discussion during my half-time seminar.

Anja Andersson and **Agneta Cedefors-Blom** for being understanding and giving me more than a helping hand in the administrative jungle of research and dissertation.

Lena Otterlind for all support and covering for me during all the years we have worked together.

Radha Korsoski for stepping in for me as Head of the complicated pregnancy and delivery unit and for being a supportive friend.

All colleagues at the Department of Obstetrics and Gynaecology, Sahlgrenska University Hospital, for covering for me when I had time off for research and for all your kind and supporting words.

My mother **Elisabeth** and father **Folke** who raised me to believe in myself and that nothing is impossible.

My two sisters **Tove** and **Ylva,** for being fantastic role models and curling me through life.

My parents in law **Jan** and **Agneta**, without your help with our kids the thesis would not have been finished. I will always be grateful for all your help.

Mikael, my husband, thank you for motivating and supporting me in completing this thesis and allowing me to prove to myself I could do it. Thank you for your understanding and for your help with proofreading.

The Sahlgrenska University Hospital LUA/ALF and the Hjalmar Svensson Research Foundation for financial support in the work with this thesis.

REFERENCES

ACOG. (2009). ACOG Practice Bulletin No. 107: Induction of labor. *Obstet Gynecol, 114*(2 Pt 1), 386-397. doi:10.1097/AOG.0b013e3181b48ef5

ACOG. (2013). ACOG Committee Opinion No 579: Definition of term pregnancy. *Obstet Gynecol, 122*(5), 1139-1140. doi:10.1097/01.AOG.0000437385.88715.4a

ACOG. (2014). Executive summary: Neonatal encephalopathy and neurologic outcome, second edition. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy. *Obstet Gynecol, 123*(4), 896-901. doi:10.1097/01.AOG.0000445580.65983.d2

ACOG. (2014). Practice bulletin no. 146: Management of late-term and postterm pregnancies. *Obstet Gynecol, 124*(2 Pt 1), 390-396. doi:10.1097/01.Aog.0000452744.06088.48

ACOG. (2015). Committee Opinion No. 644: The Apgar Score. *Obstet Gynecol, 126*(4), e52-55. doi:10.1097/aog.0000000000001108

ACOG. (2017). Committee Opinion No 700: Methods for Estimating the Due Date. *Obstet Gynecol, 129*(5), e150-e154. doi:10.1097/aog.0000000000002046

ACOG. (2019). ACOG Practice Bulletin No. 205: Vaginal Birth After Cesarean Delivery. *Obstet Gynecol, 133*(2), e110-e127. doi:10.1097/aog.0000000000003078

Aghideh, F. K., Mullin, P. M., Ingles, S., Ouzounian, J. G., Opper, N., Wilson, M. L., . . . Lee, R. H. (2014). A comparison of obstetrical outcomes with labor induction agents used at term. *J Matern Fetal Neonatal Med, 27*(6), 592-596. doi:10.3109/14767058.2013.831066

Alfirevic, Z. & Walkinshaw, SA. (1995). A randomised controlled trial of simple compared with complex antenatal fetal monitoring after 42 weeks of gestation. *Bjog*, 1995;102:638-43. doi: 10.1111/j.1471-0528.1995.tb11402.x

Alfirevic, Z., Stampalija, T., & Dowswell, T. (2017). Fetal and umbilical Doppler ultrasound in high-risk pregnancies. *Cochrane Database Syst Rev, 6*(6), Cd007529. doi:10.1002/14651858.CD007529.pub4

Alfirevic, Z., Stampalija, T., & Medley, N. (2015). Fetal and umbilical Doppler ultrasound in normal pregnancy. *Cochrane Database Syst Rev, 2015*(4), Cd001450. doi:10.1002/14651858.CD001450.pub4

Alkmark, Berglin, Dencker, Elden, Hagberg, Karlsson, . . . Jivegård. (2020). Induction of labour at 41 or 42 weeks of gestation [Igångsättning av förlossning vid 41 eller 42 fullgångna graviditetsveckor]. . *Göteborg: Västra Götalandsregionen, Sahlgrenska Universitetssjukhuset, HTA-centrum: 2020., Regional activity based HTA 2020:111*.

Alkmark, M., Keulen, J. K. J., Kortekaas, J. C., Bergh, C., van Dillen, J., Duijnhoven, R. G., . . . de Miranda, E. (2020). Induction of labour at 41 weeks or expectant management until 42 weeks: A systematic review and an individual participant data meta-analysis of randomised trials. *PLoS Med, 17*(12), e1003436. doi:10.1371/journal.pmed.1003436

Amer-Wåhlin, I., Hellsten, C., Norén, H., Hagberg, H., Herbst, A., Kjellmer, I., . . . Marsál, K. (2001). Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. *Lancet, 358*(9281), 534-538. doi:10.1016/s0140-6736(01)05703-8

Aminu, M., Bar-Zeev, S., & van den Broek, N. (2017). Cause of and factors associated with stillbirth: a systematic review of classification systems. *Acta Obstet Gynecol Scand, 96*(5), 519-528. doi:10.1111/aogs.13126

Apgar, V. (1953). A proposal for a new method of evaluation of the newborn infant. *Curr Res Anesth Analg, 32*(4), 260-267.

Arrowsmith, S., Wray, S., & Quenby, S. (2011). Maternal obesity and labour complications following induction of labour in prolonged pregnancy. *Bjog, 118*(5), 578-588. doi:10.1111/j.1471-0528.2010.02889.x

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res, 46*(3), 399-424. doi:10.1080/00273171.2011.568786

Badawi, N., Kurinczuk, J. J., Keogh, J. M., Alessandri, L. M., O'Sullivan, F., Burton, P. R., . . . Stanley, F. J. (1998). Antepartum risk factors for newborn encephalopathy: the Western Australian case-control study. *Bmj, 317*(7172), 1549-1553. doi:10.1136/bmj.317.7172.1549

Bassler, D., Briel, M., Montori, V. M., Lane, M., Glasziou, P., Zhou, Q., . . . Ramsay, T. (2010). Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *Jama, 303*(12), 1180-1187. doi:10.1001/jama.2010.310

Bassler, D., Montori, V. M., Briel, M., Glasziou, P., Walter, S. D., Ramsay, T., & Guyatt, G. (2013). Reflections on meta-analyses involving trials stopped early for benefit: is there a problem and if so, what is it? *Stat Methods Med Res, 22*(2), 159-168. doi:10.1177/0962280211432211

Berglin L, Berter B, Carlsson Y, Elden H, Karlsson EK, Sangskär H, Sengpiel V, Sjögren P, Svanberg T, Strandell A. (2020). Safety and effectiveness of outpatient compared with inpatient induction of labour in singleton uncomplicated pregnancies [Säkerhet och effektivitet vid heminduktion jämfört med igångsättning av förlossning på sjukhus]. *Göteborg: Västra Götalandsregionen, Sahlgrenska Universitetssjukhuset, HTA-centrum: 2020., Regional activity based HTA 2020:117*.

Bergsjo, P., Huang, G. D., Yu, S. Q., Gao, Z. Z., & Bakketeig, L. S. (1989). Comparison of induced versus non-induced labor in post-term pregnancy. A randomized prospective study. *Acta Obstet Gynecol Scand, 68*(8), 683-687. doi:10.3109/00016348909006139

Betrán, A. P., Ye, J., Moller, A. B., Zhang, J., Gülmezoglu, A. M., & Torloni, M. R. (2016). The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. *PLoS One, 11*(2), e0148343. doi:10.1371/journal.pone.0148343

Bhatia, M., Mitsi, V., Court, L., Thampi, P., El-Nasharty, M., Hesham, S., . . . Impey, L. (2019). The outcomes of pregnancies with reduced fetal movements: A retrospective cohort study. *Acta Obstet Gynecol Scand, 98*(11), 1450-1454. doi:10.1111/aogs.13671

Bishop, E. H. (1964). PELVIC SCORING FOR ELECTIVE INDUCTION. *Obstet Gynecol, 24*, 266-268.

Blettner, M., Sauerbrei, W., Schlehofer, B., Scheuchenpflug, T., & Friedenreich, C. (1999). Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol, 28*(1), 1-9. doi:10.1093/ije/28.1.1

Boivin, A., Richards, T., Forsythe, L., Grégoire, A., L'Espérance, A., Abelson, J., & Carman, K. L. (2018). Evaluating patient and public involvement in research. *Bmj, 363*, k5147. doi:10.1136/bmj.k5147

Campbell, S. (1969). The prediction of fetal maturity by ultrasonic measurement of the biparietal diameter. *J Obstet Gynaecol Br Commonw, 76*(7), 603-609. doi:10.1111/j.1471-0528.1969.tb06146.x

Casey, B. M., McIntire, D. D., & Leveno, K. J. (2001). The continuing value of the Apgar score for the assessment of newborn infants. *N Engl J Med, 344*(7), 467-471. doi:10.1056/nejm200102153440701

Caughey, A. B., Nicholson, J. M., Cheng, Y. W., Lyell, D. J., & Washington, A. E. (2006). Induction of labor and cesarean delivery by gestational age. *Am J Obstet Gynecol, 195*(3), 700-705. doi:10.1016/j.ajog.2006.07.003

Caughey, A. B., Stotland, N. E., Washington, A. E., & Escobar, G. J. (2007). Maternal and obstetric complications of pregnancy are associated with increasing gestational age at term. *Am J Obstet Gynecol, 196*(2), 155.e151-156. doi:10.1016/j.ajog.2006.08.040

Chen, H. Y., Grobman, W. A., Blackwell, S. C., & Chauhan, S. P. (2019). Neonatal and Maternal Adverse Outcomes Among Low-Risk Parous Women at 39-41 Weeks of Gestation. *Obstet Gynecol, 134*(2), 288-294. doi:10.1097/aog.0000000000003372

Cheng, Y. W., Kaimal, A. J., Snowden, J. M., Nicholson, J. M., & Caughey, A. B. (2012). Induction of labor compared to expectant management in low-risk women and associated perinatal outcomes. *Am J Obstet Gynecol, 207*(6), 502.e501-508. doi:10.1016/j.ajog.2012.09.019

Cnattingius, S., & Taube, A. (1998). Stillbirths and rate of neonatal deaths in 76,761 postterm pregnancies in Sweden, 1982-1991; a register study. *Acta Obstet Gynecol Scand, 77*(5), 582-583. doi:10.1034/j.1600-0412.1998.770520.x

Collins, R., Bowman, L., Landray, M., & Peto, R. (2020). The Magic of Randomization versus the Myth of Real-World Evidence. *N Engl J Med, 382*(7), 674-678. doi:10.1056/NEJMsb1901642

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med, 342*(25), 1887-1892. doi:10.1056/nejm200006223422507

Crocker, J. C., Ricci-Cabello, I., Parker, A., Hirst, J. A., Chant, A., Petit-Zeman, S., . . . Rees, S. (2018). Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. *Bmj, 363*, k4738. doi:10.1136/bmj.k4738

Crowther, C. A., Haslam, R. R., Hiller, J. E., Doyle, L. W., & Robinson, J. S. (2006). Neonatal respiratory distress syndrome after repeat exposure to antenatal corticosteroids: a randomised controlled trial. *Lancet, 367*(9526), 1913-1919. doi:10.1016/s0140-6736(06)68846-6

da Silva Charvalho, P., Hansson Bittár, M., & Vladic Stjernholm, Y. (2019). Indications for increase in caesarean delivery. *Reprod Health, 16*(1), 72. doi:10.1186/s12978-019-0723-8

de Vaan, M. D. T., ten Eikelder, M. L. G., Jozwiak, M., Palmer, K. R., Davies-Tuck, M., Bloemenkamp, K. W. M., . . . Boulvain, M. (2019). Mechanical methods for induction of labour. *Cochrane Database of Systematic Reviews*(10). doi:10.1002/14651858.CD001233.pub3

Delaney, M., & Roggensack, A. (2008). Guidelines for the management of pregnancy at 41+0 to 42+0 weeks. *J Obstet Gynaecol Can, 30*(9), 800-810. doi:10.1016/s1701-2163(16)32945-0

Dencker, A., Bergqvist, L., Berg, M., Greenbrook, J. T. V., Nilsson, C., & Lundgren, I. (2020). Measuring women's experiences of decision-making and aspects of midwifery support: a confirmatory factor analysis of the revised Childbirth Experience Questionnaire. *BMC Pregnancy Childbirth, 20*(1), 199. doi:10.1186/s12884-020-02869-0

Dencker, A., Taft, C., Bergqvist, L., Lilja, H., & Berg, M. (2010). Childbirth experience questionnaire (CEQ): development and evaluation of a multidimensional instrument. *BMC Pregnancy Childbirth, 10*, 81. doi:10.1186/1471-2393-10-81

Divon, M. Y., Ferber, A., Nisell, H., & Westgren, M. (2002). Male gender predisposes to prolongation of pregnancy. *Am J Obstet Gynecol, 187*(4), 1081-1083. doi:10.1067/mob.2002.126645

Divon, M. Y., Haglund, B., Nisell, H., Otterblad, P. O., & Westgren, M. (1998). Fetal and neonatal mortality in the postterm pregnancy: the impact of gestational age and fetal growth restriction. *Am J Obstet Gynecol, 178*(4), 726-731. doi:10.1016/s0002-9378(98)70482-x

Dodd, J. M., Crowther, C. A., Grivell, R. M., & Deussen, A. R. (2017). Elective repeat caesarean section versus induction of labour for

women with a previous caesarean birth. *Cochrane Database Syst Rev, 7*(7), Cd004906. doi:10.1002/14651858.CD004906.pub5

Dodds, S. D., & Wolfe, S. W. (2000). Perinatal brachial plexus palsy. *Curr Opin Pediatr, 12*(1), 40-47. doi:10.1097/00008480-200002000-00009

Donald, I., & Brown, T. G. (1961). Demonstration of tissue interfaces within the body by ultrasonic echo sounding. *Br J Radiol, 34*, 539-546. doi:10.1259/0007-1285-34-405-539

Dos Santos, F., Drymiotou, S., Antequera Martin, A., Mol, B. W., Gale, C., Devane, D., . . . Thangaratinam, S. (2018). Development of a core outcome set for trials on induction of labour: an international multistakeholder Delphi study. *Bjog, 125*(13), 1673-1680. doi:10.1111/1471-0528.15397

Edwards, R. K., & Richards, D. S. (2000). Preinduction cervical assessment. *Clin Obstet Gynecol, 43*(3), 440-446. doi:10.1097/00003081-200009000-00004

Einerson, B. D., Nelson, R. E., Sandoval, G., Esplin, M. S., Branch, D. W., Metz, T. D., . . . Varner, M. (2020). Cost of Elective Labor Induction Compared With Expectant Management in Nulliparous Women. *Obstet Gynecol, 136*(1), 19-25. doi:10.1097/aog.0000000000003930

El Marroun, H., Zeegers, M., Steegers, E. A., van der Ende, J., Schenk, J. J., Hofman, A., . . . Tiemeier, H. (2012). Post-term birth and the risk of behavioural and emotional problems in early childhood. *Int J Epidemiol, 41*(3), 773-781. doi:10.1093/ije/dys043

Elze, M. C., Gregson, J., Baber, U., Williamson, E., Sartori, S., Mehran, R., . . . Pocock, S. J. (2017). Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies. *J Am Coll Cardiol, 69*(3), 345-357. doi:10.1016/j.jacc.2016.10.060

Evertson, L. R., Gauthier, R. J., Schifrin, B. S., & Paul, R. H. (1979). Antepartum fetal heart rate testing. I. Evolution of the nonstress test. *Am J Obstet Gynecol, 133*(1), 29-33. doi:10.1016/0002-9378(79)90406-x

Falk, M., Nelson, M., & Blomberg, M. (2019). The impact of obstetric interventions and complications on women's satisfaction with childbirth a population based cohort study including 16,000 women. *BMC Pregnancy Childbirth, 19*(1), 494. doi:10.1186/s12884-019-2633-8

Fanshawe, T. R., & Perera, R. (2019). Conducting one-stage IPD meta-analysis: which approach should I choose? *BMJ Evid Based Med, 24*(5), 190. doi:10.1136/bmjebm-2018-111127

Finucane, E. M., Murphy, D. J., Biesty, L. M., Gyte, G. M., Cotter, A. M., Ryan, E. M., . . . Devane, D. (2020). Membrane sweeping for induction of labour. *Cochrane Database Syst Rev, 2*(2), Cd000451. doi:10.1002/14651858.CD000451.pub3

Flenady, V., Koopmans, L., Middleton, P., Frøen, J. F., Smith, G. C., Gibbons, K., . . . Ezzati, M. (2011). Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *Lancet, 377*(9774), 1331-1340. doi:10.1016/s0140-6736(10)62233-7

Franklin, J. M., & Schneeweiss, S. (2017). When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? *Clin Pharmacol Ther, 102*(6), 924-933. doi:10.1002/cpt.857

Galson, S. (2016). Real-World Evidence to Guide the Approval and Use of New Treatments. Retrieved from https://nam.edu/wp-content/uploads/2016/10/Real-World-Evidence-to-Guide-the-Approval-and-Use-of-New-Treatments.pdf

Geirsson, R. T., & Busby-Earle, R. M. (1991). Certain dates may not provide a reliable estimate of gestational age. *Br J Obstet Gynaecol, 98*(1), 108-109. doi:10.1111/j.1471-0528.1991.tb10323.x

Gelisen, O., Caliskan, E., Dilbaz, S., Ozdas, E., Dilbaz, B., Ozdas, E., & Haberal, A. (2005). Induction of labor with three different techniques at 41 weeks of gestation or spontaneous follow-up until 42 weeks in women with definitely unfavorable cervical scores. *Eur J Obstet Gynecol Reprod Biol, 120*(2), 164-169. doi:10.1016/j.ejogrb.2004.08.013

Goeree, R., Hannah, M., & Hewson, S. (1995). Cost-effectiveness of induction of labour versus serial antenatal monitoring in the Canadian Multicentre Postterm Pregnancy Trial. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne, 152*(9), 1445-1450.

Goeree, R., Hannah, M., & Hewson, S. (1995). Cost-effectiveness of induction of labour versus serial antenatal monitoring in the Canadian Multicentre Postterm Pregnancy Trial. *Cmaj, 152*(9), 1445-1450.

Golder, S., Loke, Y. K., & Bland, M. (2011). Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLoS Med, 8*(5), e1001026. doi:10.1371/journal.pmed.1001026

Goodare, H., & Smith, R. (1995). The rights of patients in research. *Bmj, 310*(6990), 1277-1278. doi:10.1136/bmj.310.6990.1277

Goodman, S., Berry, D., & Wittes, J. (2010). Bias and trials stopped early for benefit. *Jama, 304*(2), 157; author reply 158-159. doi:10.1001/jama.2010.931

Goonewardene, M., Kumara, D., Ziard, M., & Bhabu, B. (2014). Intra Cervical Foley Catheter vs oral misoprostol for pre induction cervical ripening of postdated pregnancies. *Sri Lanka Journal of Obstetrics and Gynaecology, 36*. doi:10.4038/sljog.v36i3.7714

Gradinger, F., Britten, N., Wyatt, K., Froggatt, K., Gibson, A., Jacoby, A., . . . Popay, J. (2015). Values associated with public involvement in health and social care research: a narrative review. *Health Expect, 18*(5), 661-675. doi:10.1111/hex.12158

Greenwood. (2014). Can big data tell us what clinical trials don't. *The New York Times*. Retrieved from http://www.nytimes.com/2014/10/05/magazine/can-big-data-tell-us-what-clinicaltrials-dont.html

Grivell, R. M., Alfirevic, Z., Gyte, G. M., & Devane, D. (2015). Antenatal cardiotocography for fetal assessment. *Cochrane Database Syst Rev, 2015*(9), Cd007863. doi:10.1002/14651858.CD007863.pub4

Grobman, W. A., Rice, M. M., Reddy, U. M., Tita, A. T. N., Silver, R. M., Mallett, G., . . . Macones, G. A. (2018). Labor Induction versus Expectant Management in Low-Risk Nulliparous Women. *N Engl J Med, 379*(6), 513-523. doi:10.1056/NEJMoa1800566

Grunewald, C., Hakansson, S., Saltvedt, S., & Kallen, K. (2011). Significant effects on neonatal morbidity and mortality after regional change in management of post-term pregnancy. *Acta Obstet Gynecol Scand, 90*(1), 26-32. doi:http://dx.doi.org/10.1111/j.1600-0412.2010.01019.x

Gulmezoglu, A. M., Crowther, C. A., & Middleton, P. (2006). Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev*(4), Cd004945. doi:10.1002/14651858.CD004945.pub2

Gulmezoglu, A. M., Crowther, C. A., Middleton, P., & Heatley, E. (2012). Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev*(6), Cd004945. doi:10.1002/14651858.CD004945.pub3

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., . . . Schünemann, H. J. (2011). GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol, 64*(4), 383-394. doi:10.1016/j.jclinepi.2010.04.026

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., . . . Group, G. W. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj, 336*(7650), 924-926. doi:10.1136/bmj.39489.470347.AD

Hannah, M. E., Hannah, W. J., Hellmann, J., Hewson, S., Milner, R., & Willan, A. (1992). Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. A randomized controlled trial. The Canadian Multicenter Post-term Pregnancy Trial Group. *N Engl J Med, 326*(24), 1587-1592. doi:10.1056/nejm199206113262402

Harrell, F. E., Jr., Lee, K. L., Califf, R. M., Pryor, D. B., & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Stat Med, 3*(2), 143-152. doi:10.1002/sim.4780030207

Haukoos, J. S., & Lewis, R. J. (2015). The Propensity Score. *Jama, 314*(15), 1637-1638. doi:10.1001/jama.2015.13480

Heimstad, R., Romundstad, P. R., Hyett, J., Mattsson, L. A., & Salvesen, K. A. (2007). Women's experiences and attitudes towards expectant management and induction of labor for post-term pregnancy. *Acta Obstet Gynecol Scand, 86*(8), 950-956. doi:10.1080/00016340701416929

Heimstad, R., Skogvoll, E., Mattsson, L. A., Johansen, O. J., Eik-Nes, S. H., & Salvesen, K. A. (2007). Induction of labor or serial antenatal fetal monitoring in postterm pregnancy: a randomized controlled trial. *Obstet Gynecol, 109*(3), 609-617. doi:10.1097/01.Aog.0000255665.77009.94

Henrichs, J., Verfaille, V., Jellema, P., Viester, L., Pajkrt, E., Wilschut, J., . . . de Jonge, A. (2019). Effectiveness of routine third trimester ultrasonography to reduce adverse perinatal outcomes in low risk pregnancy (the IRIS study): nationwide, pragmatic, multicentre, stepped wedge cluster randomised trial. *Bmj, 367*, l5517. doi:10.1136/bmj.l5517

Hernán, M. A., & Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol, 183*(8), 758-764. doi:10.1093/aje/kwv254

Higgins, J. P., Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj, 343*, d5928. doi:10.1136/bmj.d5928

Hildingsson, I., Karlström, A., & Nystedt, A. (2011). Women's experiences of induction of labour--findings from a Swedish regional study. *Aust N Z J Obstet Gynaecol, 51*(2), 151-157. doi:10.1111/j.1479-828X.2010.01262.x

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., . . . Loder, E. (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Bmj, 346*, f1049. doi:10.1136/bmj.f1049

Hussain, A. A., Yakoob, M. Y., Imdad, A., & Bhutta, Z. A. (2011). Elective induction for pregnancies at or beyond 41 weeks of gestation and its impact on stillbirths: a systematic review with meta-analysis. *BMC Public Health, 11 Suppl 3*, S5. doi:10.1186/1471-2458-11-s3-s5

Ingemarsson, I., & Kallen, K. (1997). Stillbirths and rate of neonatal deaths in 76,761 postterm pregnancies in Sweden, 1982-1991: a register

study. *Acta Obstet Gynecol Scand, 76*(7), 658-662. doi:10.3109/00016349709024606

Ioannidis, J. P., Haidich, A. B., Pappa, M., Pantazis, N., Kokori, S. I., Tektonidou, M. G., . . . Lau, J. (2001). Comparison of evidence of treatment effects in randomized and nonrandomized studies. *Jama, 286*(7), 821-830. doi:10.1001/jama.286.7.821

Kaimal, A. J., Little, S. E., Odibo, A. O., Stamilio, D. M., Grobman, W. A., Long, E. F., . . . Caughey, A. B. (2011). Cost-effectiveness of elective induction of labor at 41 weeks in nulliparous women. *Am J Obstet Gynecol, 204*(2), 137.e131-139. doi:10.1016/j.ajog.2010.08.012

Kelly, A. J., Kavanagh, J., & Thomas, J. (2013). Castor oil, bath and/or enema for cervical priming and induction of labour. *Cochrane Database Syst Rev, 2013*(7), Cd003099. doi:10.1002/14651858.CD003099.pub2

Kessler, J., Johnsen, S. L., Ebbing, C., Karlsen, H. O., Rasmussen, S., & Kiserud, T. (2019). Estimated date of delivery based on second trimester fetal head circumference: A population-based validation of 21 451 deliveries. *Acta Obstet Gynecol Scand, 98*(1), 101-105. doi:10.1111/aogs.13454

Keulen, J. K., Bruinsma, A., Kortekaas, J. C., van Dillen, J., Bossuyt, P. M., Oudijk, M. A., . . . de Miranda, E. (2019). Induction of labour at 41 weeks versus expectant management until 42 weeks (INDEX): multicentre, randomised non-inferiority trial. *Bmj, 364*, l344. doi:10.1136/bmj.l344

Kortekaas, J. C., Kazemier, B. M., Ravelli, A. C., de Boer, K., van Dillen, J., Mol, B., & de Miranda, E. (2015). Recurrence rate and outcome of postterm pregnancy, a national cohort study. *Eur J Obstet Gynecol Reprod Biol, 193*, 70-74. doi:10.1016/j.ejogrb.2015.05.021

Krleža-Jerić, K., Gabelica, M., Banzi, R., Martinić, M. K., Pulido, B., Mahmić-Kaknjo, M., . . . Hrgović, I. (2016). IMPACT Observatory: tracking the evolution of clinical trial data sharing and research integrity. *Biochem Med (Zagreb), 26*(3), 308-307. doi:10.11613/bm.2016.035

Kuss, O., Blettner, M., & Börgermann, J. (2016). Propensity Score: an Alternative Method of Analyzing Treatment Effects. *Dtsch Arztebl Int, 113*(35-36), 597-603. doi:10.3238/arztebl.2016.0597

Lalor, J. G., Fawole, B., Alfirevic, Z., & Devane, D. (2008). Biophysical profile for fetal assessment in high risk pregnancies. *Cochrane Database Syst Rev, 2008*(1), Cd000038. doi:10.1002/14651858.CD000038.pub2

Lawn, J. E., Blencowe, H., Pattinson, R., Cousens, S., Kumar, R., Ibiebele, I., . . . Stanton, C. (2011). Stillbirths: Where? When? Why? How to make the data count? *Lancet, 377*(9775), 1448-1463. doi:10.1016/s0140-6736(10)62187-3

Lawn, J. E., Blencowe, H., Waiswa, P., Amouzou, A., Mathers, C., Hogan, D., . . . Cousens, S. (2016). Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet, 387*(10018), 587-603. doi:10.1016/s0140-6736(15)00837-5

Lawn, J. E., Gravett, M. G., Nunes, T. M., Rubens, C. E., & Stanton, C. (2010). Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. *BMC Pregnancy Childbirth, 10 Suppl 1*(Suppl 1), S1. doi:10.1186/1471-2393-10-s1-s1

Lawn, J. E., Wilczynska-Ketende, K., & Cousens, S. N. (2006). Estimating the causes of 4 million neonatal deaths in the year 2000. *Int J Epidemiol, 35*(3), 706-718. doi:10.1093/ije/dyl043

Lawson, G. W. (2020). Naegele's rule and the length of pregnancy - A review. *Aust N Z J Obstet Gynaecol*. doi:10.1111/ajo.13253

Li, F., Wu, T., Lei, X., Zhang, H., Mao, M., & Zhang, J. (2013). The apgar score and infant mortality. *PLoS One, 8*(7), e69072. doi:10.1371/journal.pone.0069072

Lidegaard Ø , K. L., Petersen OBB, Nis Damm NP , Tabor A. (2020). Are the Danish stillbirth rates still record low? A nationwide ecological study. *BMJ Open, 10e*. doi:0.1136/bmjopen-2020-040716

Lindegren, L., Stuart, A., Herbst, A., & Kallen, K. (2017). Improved neonatal outcome after active management of prolonged pregnancies beyond 41(+2) weeks in nulliparous, but not among multiparous women. *Acta Obstet Gynecol Scand, 96*(12), 1467-1474. doi:10.1111/aogs.13237

Linder, N., Hiersch, L., Fridman, E., Klinger, G., Lubin, D., Kouadio, F., & Melamed, N. (2017). Post-term pregnancy is an independent risk factor for neonatal morbidity even in low-risk singleton pregnancies.

Arch Dis Child Fetal Neonatal Ed, 102(4), F286-f290. doi:10.1136/archdischild-2015-308553

Lindqvist, P. G., Pettersson, K., Moren, A., Kublickas, M., & Nordstrom, L. (2014). Routine ultrasound examination at 41 weeks of gestation and risk of post-term severe adverse fetal outcome: a retrospective evaluation of two units, within the same hospital, with different guidelines. *Bjog, 121*(9), 1108-1115; discussion 1116. doi:10.1111/1471-0528.12654

Lindstrom, K., Fernell, E., & Westgren, M. (2005). Developmental data in preschool children born after prolonged pregnancy. *Acta Paediatr, 94*(9), 1192-1197. doi:10.1111/j.1651-2227.2005.tb02073.x

Lou, S., Hvidman, L., Uldbjerg, N., Neumann, L., Jensen, T. F., Haben, J. G., & Carstensen, K. (2019). Women's experiences of postterm induction of labor: A systematic review of qualitative studies. *Birth, 46*(3), 400-410. doi:10.1111/birt.12412

Luthy, D. A., Malmgren, J. A., & Zingheim, R. W. (2004). Cesarean delivery after elective induction in nulliparous women: the physician effect. *Am J Obstet Gynecol, 191*(5), 1511-1515. doi:10.1016/j.ajog.2004.07.001

Mangesi, L., Hofmeyr, G. J., Smith, V., & Smyth, R. M. (2015). Fetal movement counting for assessment of fetal wellbeing. *Cochrane Database Syst Rev*(10), Cd004909. doi:10.1002/14651858.CD004909.pub3

Manning, F. A., Platt, L. D., & Sipos, L. (1980). Antepartum fetal evaluation: development of a fetal biophysical profile. *Am J Obstet Gynecol, 136*(6), 787-795. doi:10.1016/0002-9378(80)90457-3

Marsal, K., Persson, P. H., Larsen, T., Lilja, H., Selbing, A., & Sultan, B. (1996). Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatr, 85*(7), 843-848. doi:10.1111/j.1651-2227.1996.tb14164.x

Mauskopf, J. A., Paul, J. E., Grant, D. M., & Stergachis, A. (1998). The role of cost-consequence analysis in healthcare decision-making. *Pharmacoeconomics, 13*(3), 277-288. doi:10.2165/00019053-199813030-00002

MBR, S. (2020). Pregnancies, Deliveries and Newborn Infants. The Swedish Medical Birth Register. from Socialstyrelsen https://sdb.socialstyrelsen.se/if_mfr_004/val.aspx

Middleton, P., Shepherd, E., & Crowther, C. A. (2018). Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev, 5*, Cd004945. doi:10.1002/14651858.CD004945.pub4

Middleton, P., Shepherd, E., Morris, J., Crowther, C. A., & Gomersall, J. C. (2020). Induction of labour at or beyond 37 weeks' gestation. *Cochrane Database Syst Rev, 7*, Cd004945. doi:10.1002/14651858.CD004945.pub5

Mishanina, E., Rogozinska, E., Thatthi, T., Uddin-Khan, R., Khan, K. S., & Meads, C. (2014). Use of labour induction and risk of cesarean delivery: a systematic review and meta-analysis. *Cmaj, 186*(9), 665-673. doi:10.1503/cmaj.130925

Morken, N. H., Melve, K. K., & Skjaerven, R. (2011). Recurrence of prolonged and post-term gestational age across generations: maternal and paternal contribution. *Bjog, 118*(13), 1630-1635. doi:10.1111/j.1471-0528.2011.03154.x

Muglu, J., Rather, H., Arroyo-Manzano, D., Bhattacharya, S., Balchin, I., Khalil, A., . . . Thangaratinam, S. (2019). Risks of stillbirth and neonatal death with advancing gestation at term: A systematic review and meta-analysis of cohort studies of 15 million pregnancies. *PLoS Med, 16*(7), e1002838. doi:10.1371/journal.pmed.1002838

Mundle, S., Bracken, H., Khedikar, V., Mulik, J., Faragher, B., Easterling, T., . . . Weeks, A. D. (2017). Foley catheterisation versus oral misoprostol for induction of labour in hypertensive women in India (INFORM): a multicentre, open-label, randomised controlled trial. *Lancet, 390*(10095), 669-680. doi:10.1016/s0140-6736(17)31367-3

Myers, E. R., Blumrick, R., Christian, A. L., Santanu Datta, S., Gray, R. N., Kolimaga, J. T., . . . McCrory, D. C. (2002). Management of prolonged pregnancy. *Evid Rep Technol Assess (Summ)*(53), 1-6.

Nabhan, A. F., & Abdelmoula, Y. A. (2008). Amniotic fluid index versus single deepest vertical pocket as a screening test for preventing adverse pregnancy outcome. *Cochrane Database Syst Rev*(3), Cd006593. doi:10.1002/14651858.CD006593.pub2

Nakling, J., & Backe, B. (2006). Pregnancy risk increases from 41 weeks of gestation. *Acta Obstet Gynecol Scand, 85*(6), 663-668. doi:10.1080/00016340500543733

Nelson, K. B., & Ellenberg, J. H. (1984). Obstetric complications as risk factors for cerebral palsy or seizure disorders. *Jama, 251*(14), 1843-1848.

NICE. (2013). NICE Process and Methods Guides. In *Guide to the Methods of Technology Appraisal 2013*. London: National Institute for Health and Care Excellence (NICE). Copyright © 2013 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved. https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781.

NICE, G. (2008). *Induction of labour, Clinical Guideline 70*. Manchester. https://www.nice.org.uk/guidance/cg70/evidence/cg70-induction-of-labour-full-guideline2

Norman, J. E., Heazell, A. E. P., Rodriguez, A., Weir, C. J., Stock, S. J. E., Calderwood, C. J., . . . Whyte, S. (2018). Awareness of fetal movements and care package to reduce fetal mortality (AFFIRM): a stepped wedge, cluster-randomised trial. *Lancet, 392*(10158), 1629-1638. doi:10.1016/s0140-6736(18)31543-5

Norman, M., Källén, K., Wahlström, E., & Håkansson, S. (2019). The Swedish Neonatal Quality Register - contents, completeness and validity. *Acta Paediatr, 108*(8), 1411-1418. doi:10.1111/apa.14823

Norwitz, E. R., Snegovskikh, V. V., & Caughey, A. B. (2007). Prolonged pregnancy: when should we intervene? *Clin Obstet Gynecol, 50*(2), 547-557. doi:10.1097/GRF.0b013e31804c9b11

Olesen, A. W., Westergaard, J. G., & Olsen, J. (2003). Perinatal and maternal complications related to postterm delivery: a national register-based study, 1978-1993. *Am J Obstet Gynecol, 189*(1), 222-227.

Olesen, A. W., Westergaard, J. G., & Olsen, J. (2006). Prenatal risk indicators of a prolonged pregnancy. The Danish Birth Cohort 1998-2001. *Acta Obstet Gynecol Scand, 85*(11), 1338-1341. doi:10.1080/00016340600935508

Oliver, S., Bagnall, A. M., Thomas, J., Shepherd, J., Sowden, A., White, I., . . . Garrett, Z. (2010). Randomised controlled trials for policy

interventions: a review of reviews and meta-regression. *Health Technol Assess, 14*(16), 1-165, iii. doi:10.3310/hta14160

Ota, E., da Silva Lopes, K., Middleton, P., Flenady, V., Wariki, W. M., Rahman, M. O., . . . Mori, R. (2020). Antenatal interventions for preventing stillbirth, fetal loss and perinatal death: an overview of Cochrane systematic reviews. *Cochrane Database Syst Rev, 12*, Cd009599. doi:10.1002/14651858.CD009599.pub2

Pattinson, R. C., Say, L., Makin, J. D., & Bastos, M. H. (2005). Critical incident audit and feedback to improve perinatal and maternal mortality and morbidity. *Cochrane Database Syst Rev, 2005*(4), Cd002961. doi:10.1002/14651858.CD002961.pub2

Petersson, K., Persson, M., Lindkvist, M., Hammarström, M., Nilses, C., Haglund, I., . . . Mogren, I. (2014). Internal validity of the Swedish Maternal Health Care Register. *BMC Health Serv Res, 14*, 364. doi:10.1186/1472-6963-14-364

Prager, M., Eneroth-Grimfors, E., Edlund, M., & Marions, L. (2008). A randomised controlled trial of intravaginal dinoprostone, intravaginal misoprostol and transcervical balloon catheter for labour induction. *Bjog, 115*(11), 1443-1450. doi:10.1111/j.1471-0528.2008.01843.x

Pyykonen, A., Tapper, A. M., Gissler, M., Haukka, J., Petaja, J., & Lehtonen, L. (2018). Propensity score method for analyzing the effect of labor induction in prolonged pregnancy. *Acta Obstet Gynecol Scand, 97*(4), 445-453. doi:10.1111/aogs.13214

Rabin, R., & de Charro, F. (2001). EQ-5D: a measure of health status from the EuroQol Group. *Ann Med, 33*(5), 337-343. doi:10.3109/07853890109002087

Rayburn, W. F. (1987). Monitoring fetal body movement. *Clin Obstet Gynecol, 30*(4), 899-911. doi:10.1097/00003081-198712000-00014

Reddy, U. M., Bettegowda, V. R., Dias, T., Yamada-Kushnir, T., Ko, C. W., & Willinger, M. (2011). Term pregnancy: a period of heterogeneous risk for infant mortality. *Obstet Gynecol, 117*(6), 1279-1287. doi:10.1097/AOG.0b013e3182179e28

Reddy, U. M., Ko, C. W., & Willinger, M. (2006). Maternal age and the risk of stillbirth throughout pregnancy in the United States. *Am J Obstet Gynecol, 195*(3), 764-770. doi:10.1016/j.ajog.2006.06.019

Register, T. S. P. (2019). The Swedish Pregnancy Register yearly report 2018. Retrieved from https://www.medscinet.com/GR/uploads/hemsida/dokumentarkiv/2018%20GR%20%C3%A5rsrapport_2.0.pdf.

Register, T. S. P. (2020). *The Swedish Pregnancy Register yearly report 2019*. Retrieved from https://www.medscinet.com/GR/uploads/hemsida/dokumentarkiv/Graviditetsregistrets%20%C3%85rsrapport%202019_2.0.pdf

Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: rationale, conduct, and reporting. *Bmj, 340*, c221. doi:10.1136/bmj.c221

Roberts, I., & Prieto-Merino, D. (2014). Applying results from clinical trials: tranexamic acid in trauma patients. *J Intensive Care, 2*(1), 56. doi:10.1186/s40560-014-0056-1

Robinson, H. P. (1973). Sonar measurement of fetal crown-rump length as means of assessing maturity in first trimester of pregnancy. *Br Med J, 4*(5883), 28-31. doi:10.1136/bmj.4.5883.28

Robinson, H. P., & Fleming, J. E. (1975). A critical evaluation of sonar "crown-rump length" measurements. *Br J Obstet Gynaecol, 82*(9), 702-710. doi:10.1111/j.1471-0528.1975.tb00710.x

Robinson, R. (1993). Cost-benefit analysis. *Bmj, 307*(6909), 924-926. doi:10.1136/bmj.307.6909.924

Robinson, R. (1993). Costs and cost-minimisation analysis. *Bmj, 307*(6906), 726-728. doi:10.1136/bmj.307.6906.726

Roos, N., Sahlin, L., Ekman-Ordeberg, G., Kieler, H., & Stephansson, O. (2010). Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. *Acta Obstet Gynecol Scand, 89*(8), 1003-1010. doi:10.3109/00016349.2010.500009

Rosenbaum PR, R. D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rosenstein, M. G., Cheng, Y. W., Snowden, J. M., Nicholson, J. M., & Caughey, A. B. (2012). Risk of stillbirth and infant death stratified by gestational age. *Obstet Gynecol, 120*(1), 76-82. doi:10.1097/AOG.0b013e31825bd286

Rothman, K. J., Gallacher, J. E., & Hatch, E. E. (2013). Why representativeness should be avoided. *Int J Epidemiol, 42*(4), 1012-1014. doi:10.1093/ije/dys223

Rydahl, E., Declercq, E., Juhl, M., & Maimburg, R. D. (2019). Routine induction in late-term pregnancies: follow-up of a Danish induction of labour paradigm. *BMJ Open, 9*(12), e032815. doi:10.1136/bmjopen-2019-032815

Rydahl, E., Eriksen, L., & Juhl, M. (2019). Effects of induction of labor prior to post-term in low-risk pregnancies: a systematic review. *JBI Database System Rev Implement Rep, 17*(2), 170-208. doi:10.11124/jbisrir-2017-003587

SALAR. (2020). *Nationella KPP-principer version 4. Kostnad per patient* (4 ed.): The Swedish Association of Local Authorities and Regions. ISBN: 978-91-7585-881-4. https://webbutik.skr.se/bilder/artiklar/pdf/7585-881-4.pdf?issuusl=ignore

Sanchez-Ramos, L., Olivier, F., Delke, I., & Kaunitz, A. M. (2003). Labor induction versus expectant management for postterm pregnancies: a systematic review with meta-analysis. *Obstet Gynecol, 101*(6), 1312-1318. doi:10.1016/s0029-7844(03)00342-9

Sanders, G. D., Neumann, P. J., Basu, A., Brock, D. W., Feeny, D., Krahn, M., . . . Ganiats, T. G. (2016). Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *Jama, 316*(10), 1093-1103. doi:10.1001/jama.2016.12195

Sarnat, H. B., & Sarnat, M. S. (1976). Neonatal encephalopathy following fetal distress. A clinical and electroencephalographic study. *Arch Neurol, 33*(10), 696-705. doi:10.1001/archneur.1976.00500100030012

SBU. (2018). SBU Methods Resources. In *Assessment of Methods in Health Care and Social Services: A Handbook*. Stockholm: Swedish Council on Health Technology Assessment (SBU). Copyright © 2016 by the Swedish Council on Health Technology Assessment. https://www.sbu.se/contentassets/76adf07e270c48efaf67e3b560b7c59c/eng_metodboken.pdf

.

Schifrin, B. S. (1979). The rationale for antepartum fetal heart rate monitoring. *J Reprod Med, 23*(5), 213-221.

Seijmonsbergen-Schermers, A. E., van den Akker, T., Rydahl, E., Beeckman, K., Bogaerts, A., Binfa, L., . . . de Jonge, A. (2020). Variations in use of childbirth interventions in 13 high-income countries: A multinational cross-sectional study. *PLoS Med, 17*(5), e1003103. doi:10.1371/journal.pmed.1003103

Selbing, A., & Kjessler, B. (1985). Conceptual dating by ultrasonic measurement of the fetal biparietal diameter in early pregnancy. *Acta Obstet Gynecol Scand, 64*(7), 593-597. doi:10.3109/00016348509156368

Sentilhes, L., Vayssière, C., Beucher, G., Deneux-Tharaux, C., Deruelle, P., Diemunsch, P., . . . d'Ercole, C. (2013). Delivery for women with a previous cesarean: guidelines for clinical practice from the French College of Gynecologists and Obstetricians (CNGOF). *Eur J Obstet Gynecol Reprod Biol, 170*(1), 25-32. doi:10.1016/j.ejogrb.2013.05.015

SFOG-råd. (2019). Fetometri. Retrieved from www.sfog.se/media/336451/fetometri.pdf

SFOG. (2016). *SFOG Riktlinje Induktion av förlossning*. Stockholm Retrieved from https://www.sfog.se/media/301124/ind_rikt_version_klart_20160828.pdf.

Shea, K. M., Wilcox, A. J., & Little, R. E. (1998). Postterm delivery: a challenge for epidemiologic research. *Epidemiology, 9*(2), 199-204. doi:10.1097/00001648-199803000-00014

Skalkidou, A., Kieler, H., Stephansson, O., Roos, N., Cnattingius, S., & Haglund, B. (2010). Ultrasound pregnancy dating leads to biased perinatal morbidity and neonatal mortality among post-term-born girls. *Epidemiology, 21*(6), 791-796. doi:10.1097/EDE.0b013e3181f3a660

Smith, C. A., Armour, M., & Dahlen, H. G. (2017). Acupuncture or acupressure for induction of labour. *Cochrane Database Syst Rev, 10*(10), Cd002962. doi:10.1002/14651858.CD002962.pub4

Somirathne, D., Goonewardene, M., & Dasanayake, L. (2017). Three doses of oral misoprostol versus an intra-cervical Foley catheter for 24 hours for pre-induction cervical ripening in post- dated pregnancies: a randomized controlled trial. *Ceylon Med J, 62*(2), 77-82. doi:10.4038/cmj.v62i2.8470

SOS. (2018). Swedish National Board of Health and Welfare. *Dödfödda barn, en inventering och förslag på åtgärder*. Retrieved from https://stillbirth.se/wp-content/uploads/2019/01/Socialstyrelsens-rapport-2018-om-d%C3%B6df%C3%B6dda-barn..pdf

SOS. (2020). *Swedish National Board of Health and Welfare*. Staristics retrieved from https://sdb.socialstyrelsen.se/if_mfr_004/val.aspx

Sovio, U., White, I. R., Dacey, A., Pasupathy, D., & Smith, G. C. S. (2015). Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *Lancet, 386*(10008), 2089-2097. doi:10.1016/s0140-6736(15)00131-2

Stephansson, O., Petersson, K., Björk, C., Conner, P., & Wikström, A. K. (2018). The Swedish Pregnancy Register - for quality of care improvement and research. *Acta Obstet Gynecol Scand, 97*(4), 466-476. doi:10.1111/aogs.13266

Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., . . . Group, P.-I. D. (2015). Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *Jama, 313*(16), 1657-1665. doi:10.1001/jama.2015.3656

Stewart, L. A., & Tierney, J. F. (2002). To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof, 25*(1), 76-97. doi:10.1177/0163278702025001006

Stock, S. J., Ferguson, E., Duffy, A., Ford, I., Chalmers, J., & Norman, J. E. (2012). Outcomes of elective induction of labour compared with expectant management: population based study. *Bmj, 344*, e2838. doi:10.1136/bmj.e2838

Svensson, M. (2019). *Hälsoekonomisk utvärdering - Metod och tillämningar* (Vol. 1:1): Studentlitteratur. ISBN: 9789144124605

Sweden, S. (2020). Population Statistics Database. Retrieved from https://www.scb.se/en/finding-statistics/statistics-by-subject-area/population/population-composition/population-statistics/#_TablesintheStatisticalDatabase

Takikawa, K. M., Kikuchi, A., Yokoyama, A., Ono, K., Iwasawa, Y., Sunagawa, S., . . . Nakamura, T. (2008). Perinatal findings of Seckel syndrome: a case report of a fetus showing primordial dwarfism and severe microcephaly. *Fetal Diagn Ther, 24*(4), 405-408. doi:10.1159/000167269

Ten Eikelder, M. L., Oude Rengerink, K., Jozwiak, M., de Leeuw, J. W., de Graaf, I. M., van Pampus, M. G., . . . Bloemenkamp, K. W. (2016). Induction of labour at term with oral misoprostol versus a Foley catheter (PROBAAT-II): a multicentre randomised controlled non-inferiority trial. *Lancet, 387*(10028), 1619-1628. doi:10.1016/s0140-6736(16)00084-2

Tierney JF, Stewart LA, Clarke M. Chapter 26: Individual participant data. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

Tierney, J. F., Vale, C., Riley, R., Smith, C. T., Stewart, L., Clarke, M., & Rovers, M. (2015). Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. *PLoS Med, 12*(7), e1001855. doi:10.1371/journal.pmed.1001855

Timonen, S., Vara, P., Lokki, O., & Hirvonen, E. (1965). Duration of pregnancy. *Ann Chir Gynaecol Fenn Suppl, 141*, 1-33.

Tita, A. T., Lai, Y., Landon, M. B., Spong, C. Y., Leveno, K. J., Varner, M. W., . . . Mercer, B. M. (2011). Timing of elective repeat cesarean delivery at term and maternal perioperative outcomes. *Obstet Gynecol, 117*(2 Pt 1), 280-286. doi:10.1097/AOG.0b013e3182078115

Tunón, K., Eik-Nes, S. H., & Grøttum, P. (1996). A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15,000 examinations. *Ultrasound Obstet Gynecol, 8*(3), 178-185. doi:10.1046/j.1469-0705.1996.08030178.x

Vahratian, A., Zhang, J., Troendle, J. F., Sciscione, A. C., & Hoffman, M. K. (2005). Labor progression and risk of cesarean delivery in electively induced nulliparas. *Obstet Gynecol, 105*(4), 698-704. doi:10.1097/01.AOG.0000157436.68847.3b

Volpe, J. J. (2012). Neonatal encephalopathy: an inadequate term for hypoxic-ischemic encephalopathy. *Ann Neurol, 72*(2), 156-166. doi:10.1002/ana.23647

Walker, K. F., Bugg, G., Macpherson, M., McCormick, C., Grace, C., Wildsmith, C., . . . Thornton, J. (2016). Randomized trial of labor Induction in Women 35 Years of Age or Older. *N Engl J Med, 374(9),* 813-822. doi:10.1056/NEJMoa1509117

Walker, K. F., Dritsaki, M., Bugg, G., Macpherson, M., McCormick, C., Grace, N., . . . Thornton, J. G. (2017). Labour induction near term for women aged 35 or over: an economic evaluation. *Bjog, 124*(6), 929-934. doi:10.1111/1471-0528.14557

Wang, H., Rosner, G. L., & Goodman, S. N. (2016). Quantifying over-estimation in early stopped clinical trials and the "freezing effect" on subsequent research. *Clin Trials, 13*(6), 621-631. doi:10.1177/1740774516649595

Wennerholm U-B, F. A., Hagberg H, Ladfors L, Jivegård L, Svanberg T, Wessberg A, Bergh C. (2012). Induktion av förlossning vid 41 fullgångna till och med 42 fullgångna graviditetsveckor, en uppdatering av mini-HTA VGR 2007. https://alfresco.vgregion.se/alfresco/service/vgr/storage/node/content/workspace/SpacesStore/c34d819d-d9ae-4bea-81e6-36dee8f22fb6/2012_50%20HTA-rapport%20%C3%96verburenhet.pdf?a=false&guest=true

Wennerholm, U. B., Hagberg, H., Brorsson, B., & Bergh, C. (2009). Induction of labor versus expectant management for post-date pregnancy: is there sufficient evidence for a change in clinical practice? *Acta Obstet Gynecol Scand, 88*(1), 6-17. doi:10.1080/00016340802555948

Wennerholm, U. B., Saltvedt, S., Wessberg, A., Alkmark, M., Bergh, C., Wendel, S. B., . . . Hagberg, H. (2019). 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, 367,* l6131. doi:10.1136/bmj.l6131

Wessberg, A., Lundgren, I., & Elden, H. (2017). Being in limbo: Women's lived experiences of pregnancy at 41 weeks of gestation and beyond - A phenomenological study. *BMC Pregnancy Childbirth, 17*(1), 162. doi:10.1186/s12884-017-1342-4

Wessberg, A., Lundgren, I., & Elden, H. (2020). Late-term pregnancy: Navigating in unknown waters - A hermeneutic study. *Women Birth, 33*(3), 265-272. doi:10.1016/j.wombi.2019.03.011

West, H. M., Jozwiak, M., & Dodd, J. M. (2017). Methods of term labour induction for women with a previous caesarean section. *Cochrane Database Syst Rev, 6*(6), Cd009792. doi:10.1002/14651858.CD009792.pub3

Whitworth, M., Bricker, L., & Mullan, C. (2015). Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev*(7), Cd007058. doi:10.1002/14651858.CD007058.pub3

WHO. (1977). WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. *Acta Obstet Gynecol Scand, 56*(3), 247-253.

Wiberg-Itzel, E., Lipponer, C., Norman, M., Herbst, A., Prebensen, D., Hansson, A., . . . Nordström, L. (2008). Determination of pH or lactate in fetal scalp blood in management of intrapartum fetal distress: randomised controlled multicentre trial. *Bmj, 336*(7656), 1284-1287. doi:10.1136/bmj.39553.406991.25

Wollmann, C. L., Ahlberg, M., Petersson, G., Saltvedt, S., & Stephansson, O. (2017). Time-to-delivery and delivery outcomes comparing three methods of labor induction in 7551 nulliparous women: a population-based cohort study. *J Perinatol, 37*(11), 1197-1203. doi:10.1038/jp.2017.122

Wood, S., Cooper, S., & Ross, S. (2014). Does induction of labour increase the risk of caesarean section? A systematic review and meta-analysis of trials in women with intact membranes. *Bjog, 121*(6), 674-685; discussion 685. doi:10.1111/1471-0528.12328

Zizzo, A. R., Kirkegaard, I., Pinborg, A., & Ulbjerg, N. (2017). Decline in stillbirths and perinatal mortality after implementation of a more aggressive induction policy in post-date pregnancies: a nationwide register study. *Acta Obstet Gynecol Scand*, *96*(7), 862-867. doi:10.1111/aogs.13113