Ultrasound measurement of cervical length in the second trimester of pregnancy for prediction of preterm delivery

Pihla Kuusela

Department of Obstetrics and Gynecology
Institute of Clinical Sciences
The Sahlgrenska Academy
University of Gothenburg
Gothenburg, Sweden

UNIVERSITY OF GOTHENBURG

Ultrasound measurement of cervical length in the second trimester of pregnancy for prediction of preterm delivery

pihla.kuusela@vgregion.se

ISBN 978-91-7833-448-3 (PRINT)
ISBN 978-91-7833-449-0 (PDF)
http://hdl.handle.net/2077/59069

Printed by BrandFactory in Gothenburg, Sweden 2019

$$\mathrm{IISD} = \sqrt{\frac{\Sigma d_i^2}{2n}}$$

To Eemu, Uula, Aamos and Luukas

Abstract

Background: Short cervical length is a risk factor for preterm delivery (PTD) and screening of cervical length using transvaginal ultrasound in the second trimester is a proposed method to find women at risk

Aims: To examine the potential value of routine measurement of cervical length in a Swedish population of women with singleton pregnancies in the prediction of PTD. To estimate inter- and intra-observer agreement and reliability of cervical length measurements.

Methods: Women at routine fetal scan examinations in the second trimester were recruited to a study measuring cervical length with transvaginal ultrasound. Assessments were performed by certified midwives. Women were recruited at two Swedish centers in the PILOT study (Paper I) and at seven Swedish centers in the CERVIX study (Paper II). In the PILOT study the results of the measurements were not blinded but in the CERVIX study they were. The cervix was measured once in the PILOT study and twice, at least two weeks apart, in the CERVIX study. The isthmus was measured separately in the CERVIX study. The REPRODUCIBILITY study (Paper III) forms part of the CERVIX study and consists of two studies: the LIVE study and the CLIPS study. In the LIVE study, seven pairs of midwives assessed cervical length in between 24 and 30 women each. In the CLIPS study, 16 trained examiners (raters) measured cervical length twice at least two months apart on 93 video clips. The midwives were blinded to each other´s results and in the CLIPS study also to their own previous results.

Results: **Paper I**: In the PILOT study, cervical length was measured in 2122 women. Median cervical length at 16-23 gestational weeks (GW) was 39.0 mm and the prevalence of a short cervix (≤25 mm) was 0.5%. There was a significant association between cervical length and spontaneous PTD <34 GW.
Paper II: The CERVIX study included 11 456 women. The prevalence of endocervical length ≤25 mm was 4.0% at 18-20 GW (Cx1) and 4.4% at 21-23 GW (Cx2). Isthmus was present in 23% at Cx1 and in 9% at Cx2. The discriminative ability of endocervical length was better in women with no isthmus than in women with isthmus and better at Cx2 than at Cx1. At Cx1, to predict spontaneous PTD at <33 GW the best cut-off point for endocervical length was ≤29 mm, which had a

sensitivity of 43%, Area Under receiver operating characteristic Curve (AUC) of 0.68. The corresponding figure at Cx2 was ≤27 mm with AUC 0.76. Using the 27 mm cut-off at Cx2 identified 54% of spontaneous PTD before 33 gestational weeks with 35 false positive test results per one true positive and 449 women were screened to correctly identify one woman as being at risk.
Paper III: For the best examiner pair in the LIVE study the mean difference between the two examiners' measurements of endocervical length was 0.33 mm, the limits of agreement -4.06 to 4.72 mm, for the poorest examiner pair it was 0.73 mm and -11.7 to 13.2 mm, respectively. In the CLIPS study, the repeatability for the best rater was 3.9 mm and that of the poorest 9.6 mm (median 5.9 mm).

Conclusions: In the second trimester short cervical length is a risk factor for spontaneous PTD - the shorter the cervix the higher the risk. At this point cervical length has a moderate ability to identify women at risk, the disciminative ability being higher to predict early PTD (<33 GW) than PTD 34-37 GW. Inter-observer agreement and reliability of second trimester cervical length measurements differed substantially between examiner pairs in the LIVE study and so did intra-observer measurement error, repeatability and reliability between the examiners in the CLIPS study.

Key words: Cervical length measurement, preterm delivery, second trimester of pregnancy, prospective study, cohort study, observational study, mass screening, reproducibility of results, inter-observer variation, intra-observer variation, data accuracy, quality control

ISBN 978-91-7833-448-3 (PRINT)
ISBN 978-91-7833-449-0 (PDF)

List of papers

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

I. Kuusela P, Jacobsson B, Söderlund M, Bejlum C, Almström E, Ladfors L, Hagberg H, Wennerholm U-B. Transvaginal sonographic evaluation of cervical lengths in the second trimester of asymptomatic singleton pregnancies, and the risk of preterm delivery. *Acta Obstetrica et Gynecologica Scandinavica 2015;94:598-607*. *Permission for reprinting Paper I has been granted by the publisher*

II. Kuusela P, Jacobsson B, Fadl H, Wesström J, Lindgren P, Hagberg H, Wennerholm U-B*, Valentin L*. Second-trimester ultrasound measurement of cervical length for prediction of preterm delivery: a Swedish population-based multicentre observational study. *equal contribution. *In Manuscript.*

III. Kuusela P, Wennerholm U-B, Fadl H, Wesström J, Lindgren P, Hagberg H, Jacobsson B, Valentin L. Inter-and intra-observer agreement and reliability of second-trimester cervical length measurements with transvaginal ultrasound. *Submitted*

Contents

ABBREVIATIONS
1. INTRODUCTION 13
    1.1 Preterm delivery 13
    1.2 Cervix 23
    1.3 Prediction of preterm delivery 27
    1.4 Interventions to prevent spontaneous preterm delivery 31
    1.5 Measuring cervical length 36
    1.6 Intra-observer and inter-observer reproducibility and reliability 43
2. AIM 49
3. PATIENTS AND METHODS 51
    3.1 Study populations 51
    3.2 Transvaginal ultrasound 53
    3.3 National registers 55
    3.4 Overview of methods in Paper I-III 57
    3.5 Ethical permissions and comments 58
4. RESULTS AND COMMENTS 61
    4.1 Paper I 61
    4.2 Paper II 63
    4.3 Paper III 70
5. GENERAL DISCUSSION 77
    5.1 Study design 77
    5.2 Methodological aspects 78
    5.3 Strengths 80
    5.4 Limitations 81
    5.5 Differences in prevalence of short cervix 81
    5.6 Screening or not 85
6. CONCLUSION 87
7. FUTURE PERSPECTIVES 89
ACKNOWLEDGEMENTS
SAMMANFATTNING PÅ SVENSKA
REFERENCES

List of abbrevations

AUC	Area under Receiver Operating Characteristic curve
BMI	Body mass index
GW	Gestational week/s
CI	Confidence interval
Cx1	Cervical measurement between 18+0 and 20+6 gestational weeks
Cx2	Cervical measurement between 21+0 and 23+6 gestational weeks
eCRF	Electronic Case Report Form
ICC	Intraclass correlation coefficient
ICD	International Statistical Classification of Diseases and Related Health Problems
IISD	Intra-individual standard deviation
IVF	In vitro fertilization
LoA	Limits of Agreement
LR+	Positive Likelihood Ratio
LR-	Negative Likelihood Ratio
MBR	The Swedish Medical Birth Register
NNH	Number Needed to Harm
NNS	Number Needed to Screen
NPR	The Swedish National Patient Register
NPV	Negative predictive value
PPROM	Preterm premature rupture of membranes
PPV	Positive predictive value
PTD	Preterm delivery
PDR	The Swedish Prescribed Drug Register
PR	The Swedish Pregnancy Register
OR	Odds ratio
RR	Relative risk
ROC	Receiver Operating Characteristic curve
RCT	Randomized Controlled Trial
SD	Standard deviation
sPTD	Spontaneous preterm delivery
WHO	World Health Organization

1. Introduction

1.1 Preterm delivery

Definition

Preterm delivery (PTD) is defined by the World Health Organization (WHO) as delivery before 37 gestational weeks (GW) or fewer than 259 days from the first day of the woman's last menstruation[1]. PTD includes singleton and multiple deliveries resulting in a live newborn or a stillborn infant[2].

For PTD, the International Classification of Disease (ICD) encourages the inclusion of all live births. WHO's definition of stillbirth is delivery of a fetus with birth weight >500 g without evidence of life after birth[1]. ICD defines stillbirth as a stillbirth after 22+0 GW, and many countries use this cut-off to define the lower limit of PTD as ≥22+0 GW. In high income regions stillbirths contribute to 5% to10% of all PTD[2]. In Sweden, the perinatal statistics include all live births and stillbirths ≥ 22+0 GW since 2008. Before 2008, stillbirths were defined as delivery ≥28+0 GW. Miscarriages before <22+0 GW are not registered in perinatal statistics in Sweden.

PTD can be further subdivided according to the gestational age: extreme preterm (<28+0 GW), very preterm (28+0-31+6 GW) and moderate preterm (32+0-36+6 GW). Late preterm (34+0- 36+6 GW) is sometimes distinguished from moderate PTD[3]. In Sweden, between 1991 and 2001, the proportions of all PTDs were: extreme preterm 4.3%, very preterm 10.4% and moderate preterm 85.3%[4].

PTD can be divided into two categories on the basis of the clinical presentation and the onset of delivery: medically indicated PTD and spontaneous PTD (sPTD). Indicated PTD is delivery for maternal or fetal indications (a decision of the obstetrician) while sPTD starts either with spontaneous preterm prelabor rupture of the membranes (PPROM) or spontaneous preterm labour with intact membranes. Of all PTDs, about 60% to75% are spontaneous and 25% to 40% are indicated [4,5]. In 2005 in the United States, among singleton PTDs, 69% were sPTD and 31% were indicated deliveries, whereas among twin pregnancies, the rates were 45 % and 55%, respectively[6]. Etiology, treatment of symptoms and prevention of PTD (if possible) are different for indicated and spontaneous PTD. Spontaneous PTD is associated

with subclinical inflammatory or infectious processes in the membranes and placenta, and cervical collagen ripening[7,8].

Gestational age can be estimated in three ways: according to the first day of the last menstrual period, according to the first or second trimester ultrasound examination, or in the case of assisted reproductive technology, according to the date of embryo transfer. The first trimester dating scan is performed between 11 and 14 GW and the second trimester dating scan between 14 and 21 GW. In Sweden, first trimester ultrasound dating is currently recommended[9].
The formulas of Selbing and Kjessler modified by Saltvedt et al[10,11]. are used for both first and second trimester dating. Accuracy of determination of gestational age with ultrasonography is ±5 days (2 standard deviations [SD]) if biparietal diameter (BPD) is between 22 mm and 31 mm and ± 7 days (2 SD) if BPD is 32 mm to 50 mm[10,12]. In Sweden, 97% of women have gestational age assessed by routine ultrasound scan (https://www.medscinet.com/gr/dokumentarkiv.aspx).

Epidemiology

The WHO estimates the worldwide rate of PTD in 2010 as 11.1% of all live births, which is 14.9 million babies out of 135 million live births. Of these, 1.2 million occur in high-income regions and 0.5 million (42%) of the latter group occur in the United States[3]. Rates of PTD vary from about 5% in Northern European countries to 18.1% in Malawi (Figure 1).

In the United States the PTD rate decreased between 2007 (10.44%) and 2014 (9.57%) but after that it increased to 9.93% in 2017 (Births: Final data for 2017: https://www.cdc.gov/nchs/nvss/new_nvss.htm). The increase in the total PTD is mainly due to an increase among late PTDs (34-36 GW) and, in particular, among births occurring at 36 GW.

In Sweden, the total PTD rate <37 GW has been stable at between 6.1% and 5.6% between 1985 and 2016 (MFR årsrapport 2018
www.socialstyrelsen.se/publikationer2018/2018-1-6) (Figure 2).

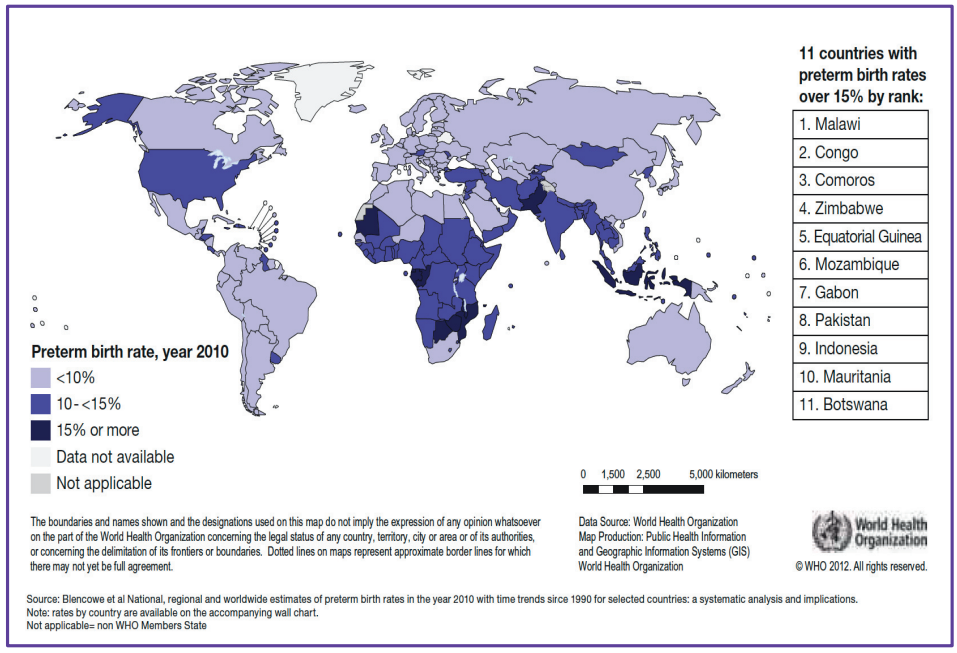


Figure 1 Preterm deliveries in 2010. (reproduced with permission from WHO. Blencowe et al 2012[3]. WHO. Born Too Soon, the Global Action Report on Preterm Birth, WHO 2012. www.who.int/maternal_child_adolescent/documents/born_too_soon/en/ Chapter 2: 15 million preterm births: priorities for action based on national, regional and global estimates, page 26, data accessed 2019-03-19)

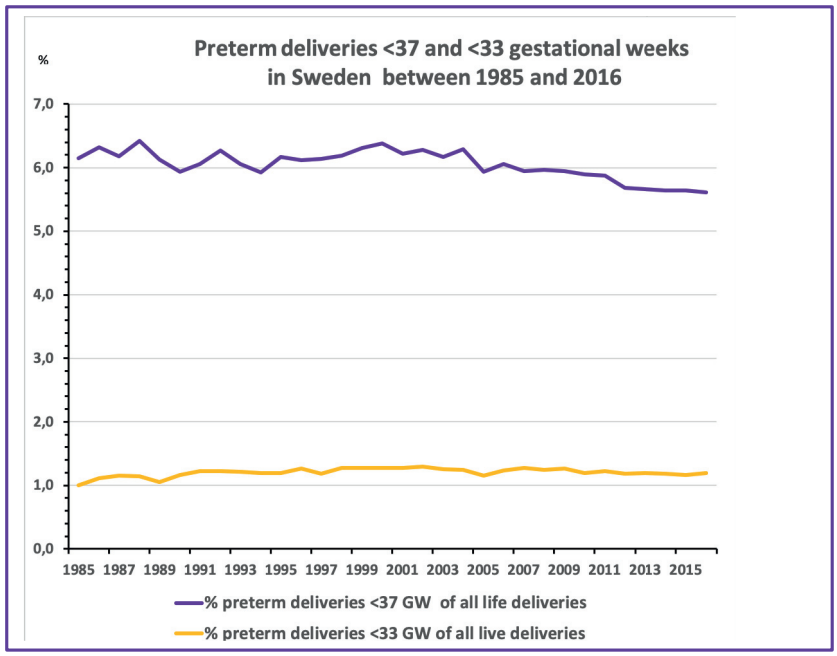


Figure 2 Preterm delivery before 37 and 33 gestational weeks in Sweden between 1985 and 2016 according data from the Swedish Medical Birth Register's annual report 2018 (www.socialstyrelsen.se/publikationer2018/2018-1-6)

Neonatal mortality, morbidity and long-term follow-up

PTD is an obstetric and neonatal problem affecting the child and its parents, the community and healthcare services. It is also associated with increased mortality, morbidity and neurological disabilities. Mortality and morbidity associated with PTD are high among extreme preterm neonates but decline with increasing gestational age[13]. Globally, the PTD is a major cause of neonatal deaths and PTD comprises 35.4% of all neonatal deaths during the first 27 days[14]. A notable proportion, 17.8%, of child deaths up to five years of age is caused by PTD complications[15]. In most high-income countries the survival rate at 24 GW is 50%, while in many low or middle-income countries the survival rate is 50% at 34 GW[2]. Interventions to prevent and treat pneumonia and diarrhoea (vaccinations, water and sanitation improvements, oral rehydration solutions), improved labour and delivery management, antenatal corticosteroids, and kangaroo mother care are the recommended methods to reduce neonatal deaths and improve the outcome for preterm babies[16,17]. A Swedish study showed a statistically significant improvement

in the survival rate of extreme preterm births (<27 GW) during recent years[18]. Two cohorts of extreme preterm infants were compared, one born between 2004 and 2007 and the other born between 2014 and 2017. One-year survival rates were 70% and 77%, respectively[18].

Necrotizing enterocolitis, pulmonary complications (respiratory distress syndrome, bronchopulmonary dysplasia, pneumothorax), retinopathy of prematurity and blindness occur more frequently in preterm than in term neonates and especially in neonates born <27 GW or with birthweight <1000g[13,19]. The onset of PTD, spontaneous or indicated, may influence neonatal morbidity due to different etiologies[20]. However, a study of extreme preterm neonates showed no difference in morbidity and mortality for spontaneous versus indicated PTD after adjustment for birth weight[21].

Neurological sequele, especially cerebral palsy (CP), are more common in infants born extremely or very preterm (Figure 3)[22]. CP causes a lifelong disability and affects about 2 per 1000 live births in Sweden[22]. For infants born <28 GW the prevalence of CP was 59 per 1000 live births in a Swedish study performed between 2007 and 2010[22]. The prevalence of CP according to gestational week at delivery, between 1973 and 2010, is shown in Figure 4[22].

Cognitive, language and motor development were assessed using the Bayley-III scale at the age of 2.5 years in infants born <27 GW (of live births 69% of infants survived to 2.5 years)[23]. Moderate or severe disability was found in 27% of infants. Among those born at 23 GW, 51% had moderate or severe disability, and among those born at 26 GW the corresponding figure was 17%. A study from Norway analysed a cohort of adults aged 19 to 35 years born preterm between 1967 and 1983, looking at long-term effects[24]. Compared with adults born at term, they found that significantly more adults born preterm had CP, mental retardation, autism, disorders of psychological development, behavior and emotions, other major disabilities (blindness, hearing loss, epilepsy) and medical disabilities severely affecting working capacity. Also income, educational level and levels of biological parenthood were lower for adults born preterm[24].

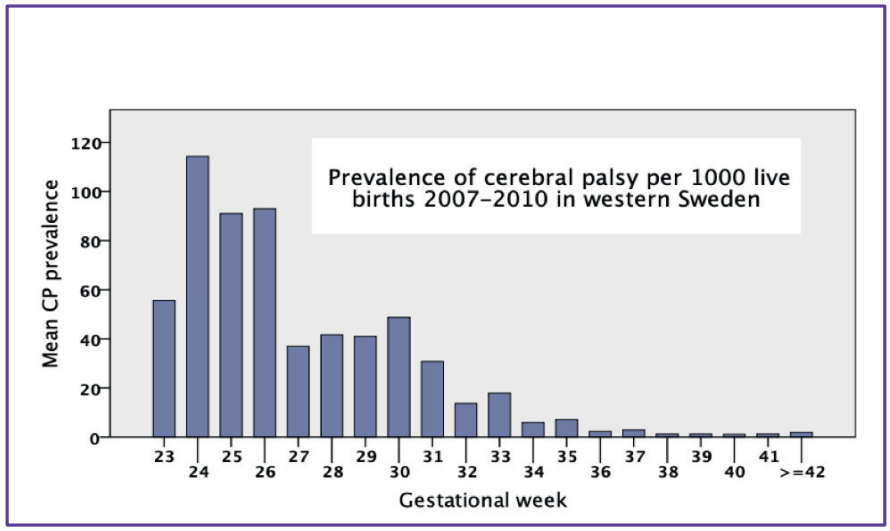


Figure 3 Prevalence of cerebral palsy per 1000 live births between 2007 and 2010 in Sweden. (with permission from Himmelman 2019, personal communication)

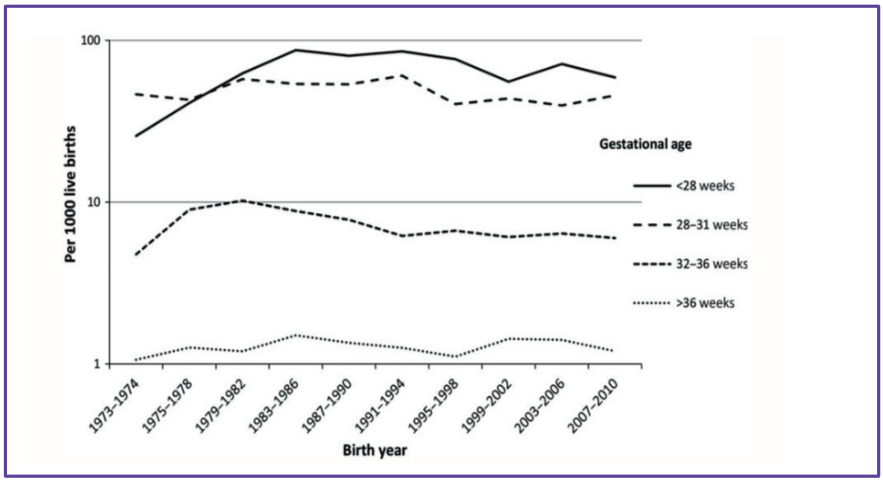


Figure 4 Prevalence of cerebral palsy by gestational age group, in birth years 1975-2010 (Himmelmann, K., & Uvebrant, P. (2018). The panorama of cerebral palsy in Sweden part XII shows that patterns changed in the birth years 2007-2010. *Acta Paediatr, 107*(3), 462-468[22], reprinted with permssion from John Wiley and Sons, Acta Pediatrica)

Risk factors

Why do some women deliver too early? This is still largely an unsolved question. Most of the women who deliver too early have no risk factors[25]. However, there are many maternal characteristics that have been associated with PTD. If a women with risk factor/s for PTD can be identified, prophylactic treatment can be targeted better.

Maternal characteristics associated with the risk of preterm delivery can be divided into factors which include reproductive history and the characteristics of the current pregnancy. The most common risk factors are listed in Table 1 and categorized into slightly increased risk (+, relative risk [RR] 1.1 – 1.6), moderate increased risk (++, RR 1.6 – 2.3) or high risk (+++, RR >2.3), adapted from[5,26-39].

Multiple pregnancy is a strong risk factor for PTD. In Sweden in 2016, 42.5% of multiple deliveries occurred before 37 GW (1438/3383), and comprised 21% of all preterm deliveries (1438/6791) (www.socialstyrelsen.se/publikationer2018/2018-1-6). Uterine overtension mediated by inflammatory cytokines and prostaglandins is thought to induce sPTD in women with multiple pregnancies[40].

Nulliparous women have higher rates of PTD than multiparous women. Nulliparous (singletons) comprised 53% of all PTD and 43% of all term deliveries in Sweden between 1992 and 2010, counted from data in Cnattingius et al. JAMA 2013[41].

The rate of PTD varies with race/ethnicity[5,31]. In United States black non-Hispanic women had a total PTD <37 GW rate of 13.9%, white non-Hispanic 9.1% and Hispanic 9.6%, respectively. (Births: Final data for 2017: https://www.cdc.gov/nchs/nvss/new_nvss.htm).

Important risk factors for PTD in singleton pregnancies are a a history of sPTD and short cervical length, as measured by transvaginal ultrasound. Short cervical length will be discussed further under "Prevalence of short cervix" in chapter 1.2 p. 26.

A history of sPTD is a strong risk factor, the earlier the previous sPTD the higher the risk. A meta-analysis of 25 studies (52,070 women) showed an absolute risk of 30% (95% confidence interval [CI] 27% – 34 %) for a recurrent sPTD in women with at least one sPTD <37 GW[36]. The risk of a recurrent sPTD after a prior PPROM <37 GW was 7% (95% CI 6% – 9%) (four studies, 3138 women), while the recurrence

risk of sPTD after a spontaneous preterm labour ≤37 GW was 23% (95% CI 13% – 33%) (three studies, 2852 women)[36].

The number of prior PTD also affects the risk. The risk of PTD <37 GW increased tenfold if woman had two prior deliveries <37 GW, compared with women with two prior term deliveries[32]. A history of prior indicated PTD increased the risk of indicated PTD between seven and eightfold in subsequent pregnancies, but also the risk of sPTD by RR 1.0 – 2.7[6,32,33]. The risk of sPTD <37 GW also depended on whether the prior PTD or the index pregnancy was a twin or singleton pregnancy[42]. The highest absolute risk, 57%, was among women with a twin gestation after a prior singleton PTD[42].

Women with a history of cervical conization have an increased risk of PTD. The risk for PTD <37 GW is almost three times higher (RR 2.61; 95% CI 1.02-3.20) for PTD <37 GW, and with repeated treatments the RR increases to fivefold (RR 5.15, 95% CI 2.45 – 7.84). Large or repeated cones increase the risk twofold (RR 2.45, 95% CI 1.38 – 4.53) when compared to smaller or medium-sized cones[43]. Weinman et al found a RR of 2.15 (95% CI 1.16-3.98) for PTD <37 GW when the cone size was ≥1.0 cm and the risk was higher among women who delivered within one year after the surgery a RR 3.26 (95% CI 1.05 – 1.93)[44].

There are different risk factors for indicated or spontaneous PTD, including whether the sPTD starts with contractions (preterm labor) or with PPROM. Assisted reproductive technology, preeclampsia and placental abruption are more likely to be associated with indicated preterm delivery than spontaneous preterm delivery [45]. Women who experience PPROM are more likely to be nulliparous, smokers, of black ethnicity or to have infections than those who experience preterm labour[34,45].

Table 1 Risk factors for preterm delivery

Risk factor	Risk for PTD or sPTD
Maternal characteristics	
Ethnicity, black	++
Smoking	++
Drug or alcohol abuse	+
Low BMI (<19 kg/m^2)	+
Depression	+
Uterine anomalies	++
History of cervical conization	++
Family history of preterm delivery, genetics	+
Low socio-economic status	+
Low (<16 years) or high (<35) maternal age	+
Intercurrent maternal disease	+
Nulliparity (no prior deliveries)	+
Reproductive history	
Prior preterm delivery	+++
Prior spontaneous preterm delivery	+++
Prior indicated delivery	+++
Prior stillbirth	++
Prior miscarriage	+
Priot induced abortion	+
Short interpregnancy interval (< 6 months)	+
Current pregnancy characteristics	
Assisted reproductive technologies (IVF, ICSI, OD)	+
Maternal complication during pregnancy (preeclampsia, diabetes)	+
Fetal growth restriction or distress	+
Fetal malformation	++
Multiple gestation	+++
Polyhydramnios	++
Infections (genitourinary or extragenital)	+
Periodontal disease	+
Stress	+
Short cervical length (≤25mm)	+++

+; RR 1.1-1.5, ++; RR 1.6-2.3, +++; RR >2.3
BMI= body mass index; ICSI= intracytoplasmic sperm injection; IVF= in vitro fertilzation pregnancy; OD= oocyte donation

Preterm delivery syndrome

The underlying mechanisms behind sPTD are many and the women who deliver preterm are a heterogenous group. The parturition mechanisms, uterine contractions, ripening and dilatation of the cervix, are the same in term and sPTD patients and follow a common process[8]. In labour at term this common pathway is activated physiologically, whereas in sPTD several disease processes activate one or more components in the process. Therefore, it is suggested that sPTD be considered as a syndrome associated with multiple mechanisms of disease[7,46]. The proposed mechanisms of disease are: microbial-induced inflammation, decidual haemorrhage and vascular disease, disruption of maternal fetal tolerance, decline in progesterone action and other mechanisms of disease (such as uterine overdistension, stress) and are shown in Figure 5[7].

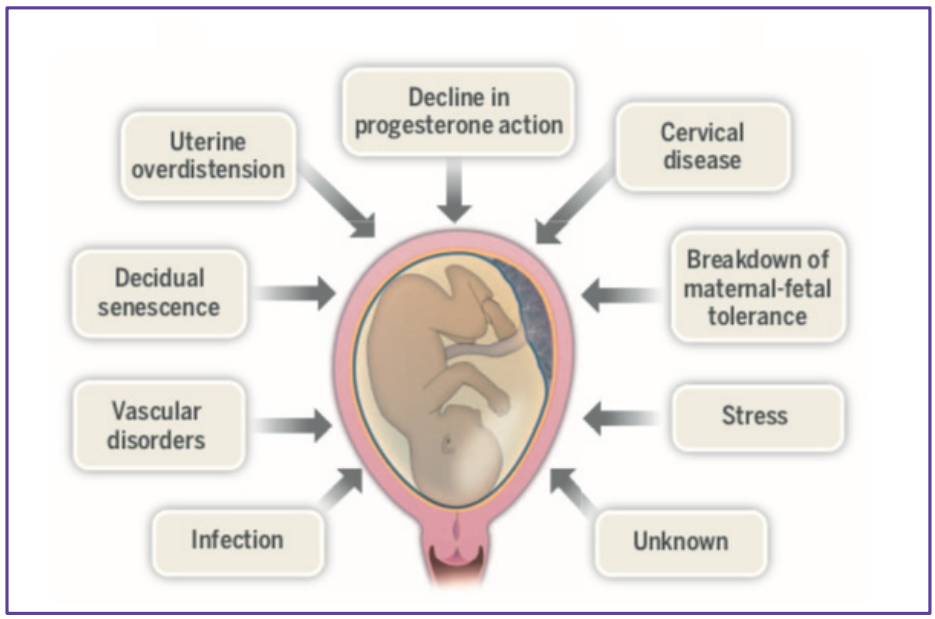


Figure 5 Proposed mechanisms of disease implicated in spontaneous preterm delivery (From: [Romero, R., Dey, S.K., & Fischer, S.J. (2014). Preterm labor: one syndrome, many causes. Science, 345(6198), 760-765][7], reprinted with permission from AAAS)

1.2 Cervix

Anatomy and histology

Cervix uteri is a cylinder which forms the lowest part of the uterus. It is situated between the corpus uteri and vagina in the vertical axis and between the bladder and rectum in the anterior-posterior axis. It bulges into the anterior vaginal wall and the visible part seen from the vagina is called the portio. The endocervical canal in the middle connects the uterine cavity to the vagina. The outer opening of the cervix to the vagina is called the external os and the inner opening to the uterus is called the internal os. The cervical upper third, called the isthmus of the cervix uteri, gradually converts to make part of the uterus during the second trimester of pregnancy, forming the lower uterine segment[47]. The lower uterine segment is fully developed at around 28 GW. The consistency of the non-pregnant cervix is quite firm.

Histologically, the endocervical canal has a 2 to 3 mm thick mucosa layer with cylinder epithelium and glands without spiral arteries. The cervix consists mainly of extracellular connective tissue including collagen, glycosaminoglycans and proteoglycans. Fibronectin and elastin are found between collagen and smooth muscle fibers. Around the external os of the cervix the external cylinder epithelium layer changes to squamous epithelium and the transformation zone can be seen on the portio, using a colposcopy. The vaginal squamous epithelium has no glands and all the lubricating mucus comes from the cervical glands[48].

Cervical softening and ripening

During pregnancy, the anatomy of the uterus and cervix change when the uterus grows, when smooth muscle cells become hypertrophic and increase seven to ten times in size. The cervix remains closed during pregnancy but undergoes changes in its muscle and connective tissue layers prior to delivery, leading to softening (Hegar sign) and effacement. During delivery the cervix shortens further and dilatates to allow the passage of the circumference of the baby's head.

Softening, ripening/remodelling, dilatation and repair postpartum are the four phases of cervical activation[8]. Together with processes in activation of chorion, amnion and myometrium this leads to parturition[8]. The first cervical changes occur directly after conception and continue until 32 GW. This is a slow process

whereby degradation of collagen makes the cervix tissue softer, which is regulated by high levels of progesterone and low levels of estradiol in the tissue[49]. The cervical ripening starts days or weeks before delivery. It is mediated by decreased progesterone production, increased progesterone metabolism and the increasing production of estradiol and relaxin[50]. When the disorganization of collagen fibrils increases and extracellular levels of glukosaminoglycans increase as does tissue hydration, this lead to increasing elasticity and shortening of cervical length[8,51]. Matrix metalloproteinases activity increases in the cervix, the chorion and the decidua, causing degradation of extracellular collagen. The process, together with plasminogen activation, results in liberation of the extracellular matrix protein, promoting the ripening, then the rupture of the fetal membranes and detachment of the placenta[52].

Changes in the extracellular matrix during ripening include an influx of inflammatory cells (macrophages, neutrophils, mast cells) into the cervical stroma and cause an inflammatory response. These cells produce cytokines and prostaglandins. Prostaglandins promote cervical ripening and artificial prostaglandin analogues are used widely for induction of delivery[39] is evidence that inflammation has a role in preterm ripening of the cervix but inflammatory cells are not necessary for term ripening of the cervix[50]. Pro-inflammatory cytokines, like interleukin -1, -6, -8, prostaglandin E2, cyclooxygenase-2 and tumor necrosis factor-α have a role in preterm cervical ripening and these mediators also increase myometrial contractility and rupture of the membranes[46].

Dynamic cervical changes during pregnancy

The cervix undergoes changes during the whole pregnancy. Longitudinal studies show that the cervix shortens gradually from the first to the third trimester[53-56]. Gramellini et al followed normal pregnancies from first trimester until term (37-40 GW)[57]. In nulliparous women there was a mean shortening of cervical length from 52 mm in the first trimester, to 45 mm in the second trimester and to 35 mm at term. The corresponding figures for multiparous women were 55 mm, 45 mm and 40 mm. Figure 6[57]. With transvaginal ultrasound examination changes in cervical length can be observed during a contraction of the uterus.

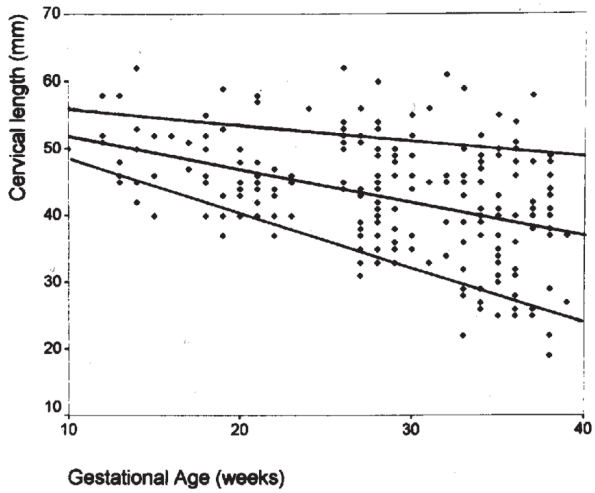


Figure 6 Cervical length in relation to gestational age in normal pregnancy. Straigth lines indicate 10th, 50th and 90th percentiles. (Gramellini D et al. Transvaginal sonographic cervical length changes during normal pregnancy. J Ultrasound Med 21: 227-232, 2002, Reprinted with permission from John Wiley and Sons[57].)

Short cervix

Maternal height has been associated with cervical length e.g. short stature – short cervix, but the study results are not conclusive[58,59]. A significant correlation between maternal weight and BMI has been shown, with the lower weight or BMI correlating with a shorter cervix[59]. White ethnicity has a significant correlation with a longer cervix [59-61]. Most of these associations are based on small differences, between 1 mm and 3 mm in mean cervical length[60].

A history of spontaneous preterm delivery, and a history of prior late miscarriage are associated with a short cervix[60].

Although a history of cervical conization is a risk factor for preterm delivery, in some studies conization was not a significant risk factor for a short cervix[60,61].

Cervical insufficiency or cervical incompetence is a weakness, shortening and asymptomatic opening of the cervix which occurs in about 1% of the total obstetric population and in 8% of the population with recurrent mid-trimester losses[62].

Prevalence of a short cervix

There is no universally agreed definition of a short cervix. The cervical length ≤25 mm is the most common cut-off used in many studies[63]. For clinical use it is convenient to have a dichotomous value to divide women into those who have a "short" and those who have a "normal" cervical length. However, a rigid cut-off does not take into account the distribution of cervical length in a population. The 10th percentile of cervical length varies from 26 mm to 36 mm in the second trimester in observational studies Table 2[55,64,65].

The prevalence of a cervical length ≤25 mm, measured in the second trimester by transvaginal ultrasound, was between 0.3% to 10 % in the largest observational studies shown in Table 2[55,64-69]. The large variation in the prevalence may depend on the study population (high-risk or low-risk, ethnicity and other background variables) or on the measurement technique, education and experience of transvaginal assessment of cervical length.

Table 2 Overview of study populations and prevalece of short cervix in observational studies

Author, journal, year	Number of screened women, race; prior PTD, country	Gestational week at screening	Cervical length	Prevalence of cervical length
Esplin et al JAMA 2017[66]	10 038 61% white 5.0% prior PTD Multicentre, USA	16-22 Median 19.0	≤15 mm ≤20 mm **≤25 mm**	1% 1.3% **2.5%**
Heath et al UOGS 1998[67]	2702 44% Caucasian 48% Afro-Caribbean 3.9 % Prior PTD	22-24 Median 23	≤15 mm ≤20 mm **≤25 mm** ≤30 mm	1.7% 3.4% 8.0% 18.0%
Iams et al NEJM 1996[64]	3000 63% black 16% prior PTD Multicenter, USA	22-24 53% at 23 45% at 24	≤20 mm **≤25mm** ≤30 mm	5% **10%** 25%
Leung UOG 2015[68]	2880 100% Chinese 6.9% prior PTD Hong Kong	18-22 Mean 20.1	≤20 mm **≤25 mm** ≤30 mm	0.2% **1.8%** 10%
Taipale and Hiilesmaa UOG 1998[69]	4206 99% white Finland	18-22 Mean 20.4	**≤25 mm** ≤29 mm ≤31 mm	**0.3%** 3% 9%
van der Ven et al AOGS 2015[65]	16 204 88% white 0% prior PTD Multicentre, The Netherlands	16-21 Median 20+2	≤15 mm **≤25 mm** ≤30 mm	0.08% **0.5%** 1.8%
Wulff et al UOG 2018[55]	3477 97,5% Caucasian 2.9% prior PTD Multicenter, Danmark	19-21 Median 19+6	≤ 15 mm ≤ 20 mm **≤ 25 mm**	0.24% 0.3% **0.78%**
PTD= preterm delivery				

1.3 Prediction of preterm delivery

Measures of diagnostic accuracy

How well can a "short cervix" (≤25 mm) discriminate between PTD <33 GW or delivery ≥33 GW? The discriminative ability can be quantified by different measures of diagnostic accuracy as shown in the example below.

Example test	Delivery <33 GW	Delivery ≥33 GW	Total
Cervix ≤25mm	*a* true positive	*b* false positive	*a+b*
Cervix >25mm	*c* false negative	*d* true negative	*c+d*
Total	*a+c*	*b+d*	*a+b+c+d*

Sensitivity = $\frac{true\ positive}{All\ Delivered<33\ GW} = \frac{a}{a+c}$

Specificity = $\frac{true\ negative}{All\ Delivered\geq 33GW} = \frac{d}{b+d}$

Positive predictive value (PPV) = $\frac{true\ positive}{All\ with\ cervix\ \leq 25\ mm} = \frac{a}{a+b}$

The PPV tell us the proportion of those with cervix ≤25 mm and who deliver <33 GW

Negative predictive value (NPV) = $\frac{true\ negative}{All\ with\ cervix\ >25\ mm} = \frac{d}{c+d}$

The NPV tell us the proportion of those with cervix >25 mm and who deliver >33 GW

Positive Likelihood ratio (LR+) = $\frac{Sensitivity}{1-Specificity}$

LR+ tells us how many times the odds of PTD <33 GW increases if the cervical length is ≤ 25 mm.

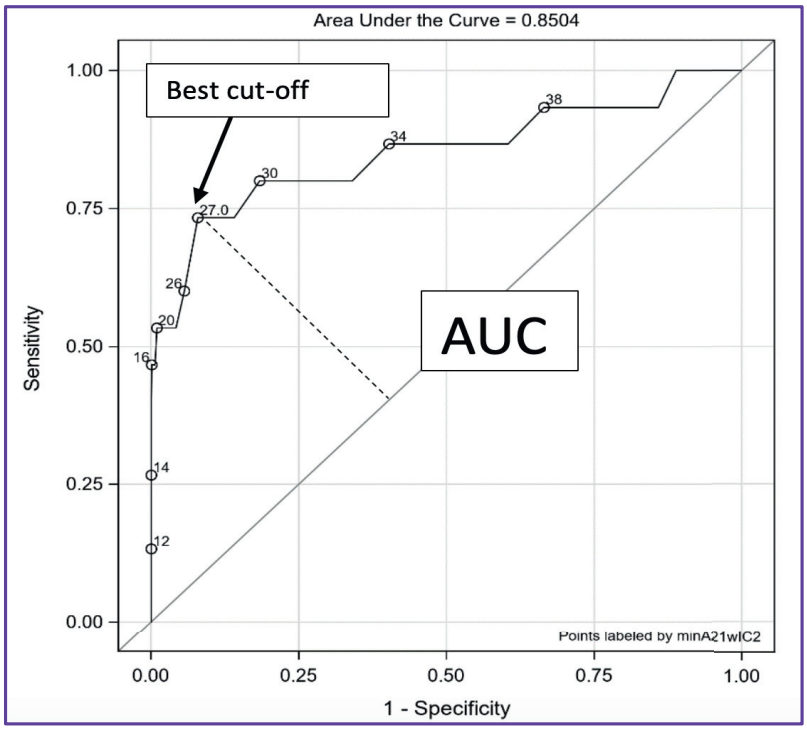
Negative Likelihood ratio (LR-) = $\frac{1-Sensitivity}{Specificity}$

LR- tells us how many times the odds of PTD <33 GW decreases if the cervical length is >25 mm.

The mathematically optimal cut-off point of cervical length for predicting preterm delivery at a particular GW can be determined from the Receiver Operating Characteristic (ROC) curve (Figure XX). It is the point on the ROC curve situated

farthest from the reference line on the ROC curve which is equal to the point at maximize the sum of sensitivity and specificity.

Area Under the Curve (AUC) can be used as a measure of predictive ability.



Number needed to harm (NNH)

The number (ratio) of false positive results per one true positive test result.

$$\text{NNH} = \frac{b}{a}$$

Number needed to screen (NNS)

Number of women needed to screen to detect one true positive test result

$$\text{NNS} = \frac{a+b+c+d}{a} = \frac{Total\ number}{a}$$

Prediction of spontaneous preterm delivery using fetal fibronectin

The main screening modalities that have been studied for prediction of sPTD are assessment of cervical length and fetal fibronectin (fFN)[70]. Cervical length assessment is studied in asymptomatic women with singleton pregnancy at high risk or low risk (universal screening), in women with multiple pregnancy and in symtomatic women with threatened labour[70,71]. Cervical length and prediction of sPTD are presented in Chapter 1.3 p. 30 and in Table 4 and Table 5 in p. 40-43.

Fetal fibronectin is a bedside test using vaginal or cervical secretions. Exceeding the qualitative test cut-off (≥ 50 ng/mL) is associated with an increased risk of sPTD. Qualitative fFN and different threshold values are also studied to predict sPTD.

Qualitative fFN and its predictive value has been studied in asymptomatic women with singleton pregnancies alone or in combination with cervical measurement, for symptomatic woman with singleton pregnancies, and for women with multiple pregnancies. For asymptomatic women, fFN was evaluated in the Preterm Prediction study, and it had a PPV of between 17% and 25% to predict sPTD <34 GW, depending on the gestational age at the time of sampling[72]. For symptomatic women with preterm labour the fFN test can be used to decide whether to hospitalize or not. A negative test result (<50 mg/ml) in symptomatic women before 34 GW had a NPV of 99.5% for no delivery within seven days[73].

Quantitative fFN tests have been evaluated for thresholds 10, 50, 200 and 500 ng/mL in symptomatic women with a PPV of 19%, 32%, 61%, and 75%, respectively, to predict sPTD within 14 days[74]. For asymptomatic high-risk women at the same thresholds, the PPV was 17%, 24%, 38% and 48% to predict sPTD<34 GW when tested in late second trimester[75]. Esplin et al evaluated self-collected vaginal fFN in nulliparous women with a singleton pregnancy in their first trimester (6 to 11 GW), in the second trimester (16 to 22 GW) and between 22 to 33 GW with three thresholds of fFN: ≤10, ≤50 and ≤200 ng/mL[66]. All gestational ages and thresholds had very low PPV, between 2% to 6% for prediction of sPTD <32 GW and between 6% to 14% for prediction of sPTD <37 GW. When they combined cervical length and fFN, AUC was 0.67 (95% CI 0.64 – 0.70) for prediction of sPTD <37 GW, fFN alone had an AUC of 0.59 (95% CI 0.56- 0.62) and cervical length alone had an AUC of 0.67 (95% CI 0.64 - 0.70) when screening was performed between 22 and 30 GW[66]. The optimal cut-off of fFN was found at 7.1 ng/mL and cervical length at 31.7 mm. Their findings do not support screening of cervical length or fFN in asymptomatic low-risk woman.

Prediction of preterm delivery using cervical length measurement

Asymptomatic women in the second trimester, with a singleton pregnancy and a short cervix have an increased risk of sPTD. Iams et al found RR 6.2 for sPTD <35 when cervical length was below 26 mm (10th percentile) and RR 9.9 at cervical length below 22 mm (5th percentile)[64]. Taipale and Hiilesmaa found RR 8 at cervical length 29 mm (3rd percentile) for sPTD <35 GW[69].

How good is cervical length at predicting PTD and which cut-off gives the best estimations? Observational studies have used such predictive estimates as sensitivity, specificity, PPV, NPV, LR+ and LR- for prediction of PTD. To predict sPTD <34 GW or <35 GW in blinded observational studies, PPV varied between 14% and 40% when cut-off of cervical length was ≤ 20 mm or ≤ 25 mm, NPV varied between 97% and 99%[25,64,67,69,76]. To predict sPTD <32 GW using cervical length measurement, the AUC varied between 0.57 to 0.64, and to predict sPTD <37 GW the AUC was 0.51 to 0.72 in low-risk populations[65,66,77]. In Chapter 1.5.5., a systematic review is presented of cervical measurement of asymptomatic women at low risk of PTD using transvaginal ultrasound in the second trimester.

Women at high risk of PTD may have more benefits from screening. Women with a history of sPTD and a short cervical length in a current pregnancy are at high risk of sPTD[78]. A cervical length <25 mm had a PPV of 20% and NPV of 97% for prediction of sPTD <34 GW in the study of 469 women by Guzman et al. To predict PTD <34 GW was AUC 0.76 and for PTD <30 GW AUC was 0.85 when cervical measurement between 15 and 24 GW was assessed in a high-risk population[78]. For prediction of sPTD <35 GW, the corresponding values are 75% (PPV) and 77% (NPV) in a small study by Owen et al of a high-risk group[78,79].

Women with multiple pregnancies are at high risk of PTD and have shorter cervical lengths than women with singleton pregnancies, while a cervical length <25 mm is an independent predictor of sPTD[80]. However, so far, its predictive value is unclear and routine screening of the population for short cervix length is not recommended[70].

1.4 Interventions to prevent preterm delivery

Primary preventive interventions to reduce the consequences of preterm delivery are directed at all women. Secondary interventions are limited to a sub-group who have known or identified risk factor/s, and tertiary interventions to women showing the first signs of early parturition[81]. A recent systematic review of 112 reviews of primary and secondary interventions found positive effects from some primary preventions: lifestyle and behavioural changes, nutritional supplements (calcium and zink included) and screening for lower genital tract infections[82]. Screening for, and treatment of, asymptomatic bacteriuria have also been shown to reduce the rate of PTD[83].

Secondary preventive interventions for high risk women shown to have a significant effect on the prevention of PTD are cerclage, progesterone, low-dose aspirin, lifestyle and behavioural changes. Aspirin reduces the risk of PTD <37 GW and PTD <34 GW in women at risk of preeclampsia, RR 0.93 (95% CI 0.89-0.98) and RR 0.90 (95% CI 0.83-0.98), respectively[84]. The systematic review by Matei et al reports positive effects from treating periodontal disease, stopping smoking, clindamycin for bacterial vaginosis, and calcium supplementation in women at risk of hypertensive disorders in some studies[82].

Interventions for preventing PTD can be considered for women with a history of late pregnancy loss or PTD. Other risk factors for PTD include uterine malformation, multiple pregnancy, conization, cervical insufficiency, short cervical length in the first or second trimester. Interventions which have been studied in women with these risk factors including short cervical length are progesterone[85-88], cervical cerclage[89] or pessary[90,91]. A recent systematic review with a meta-analysis of 40 studies (comprising 11,300 women) of progesterone, pessary and cerclage concluded that vaginal progesterone was the only intervention that could prevent preterm delivery[92].

Tertiary preventive treatments of sPTD in symptomatic women are: early diagnosis of preterm labor, short term (max 48 hours) tocolysis to delay delivery that allows transport and corticosteroid treatment (for reducing neonatal morbidity), and the detection and treatment of infections[81]. There is also the so-called rescue cerclage in emergencies, where the amniotic membranes are bulging and there is a risk of miscarriage.

Cerclage

Cervical cerclage (cervical stich) is a suture placed around the cervix with the purpose of providing mechanical support of the cervix (Figure 7). Two transvaginal techniques are used: the McDonald suture or the Shirodcar suture. Transabdominal cerclage can be considered for women who have had a trachelectomy procedure. The transvaginal suture is placed under local, spinal or general anaesthesia. A history of cervical insufficiency or a very short cervix in first or second trimester can lead to second trimester pregnancy loss or extremely preterm delivery, and cerclage can be offered to those women as prophylactic treatment. A recent systematic review showed that cerclage in women at high risk had RR of perinatal death of 0.82 (95% CI 0.65-1.04)[93]. For the risk of PTD <34 GW, the average RR was 0.77 (95% CI 0.66-0.89), compared with those without cerclage[93] (Table 3).

The risk of maternal pyrexia increases after cerclage placement, which should be taken into account[93,94]. For other complications such as PPROM and chorioamnionitis no significant risks were found.

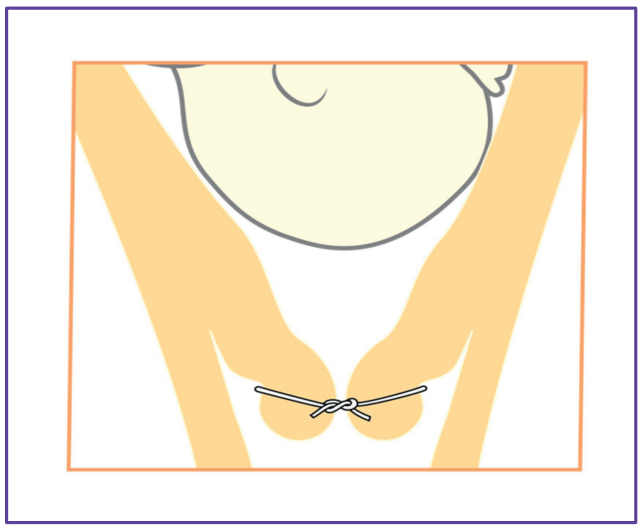


Figure 7 Cerclage placed around cervix (permission from @bo.jacobsson)

Pessary

The pessary is a silicone, concave cup-formed device with a hole in the middle. It can be applied around the cervix to keep the cervical canal closed and change the inclination of the cervical canal, thereby reducing the pressure on the internal os (Figure 8). There are conflicting results from RCTs and systematic reviews (Table 3) about the effect of a pessary in reducing the rate of PTD, neonatal morbidity and mortality in singleton pregnancies[90,91,95-97]. A single center study in women with a history of PTD and cervical length ≤25 mm showed a reduction of sPTD <34 GW by 52%, RR 0.48 (95% CI 0.24-0.95), but adverse events (vaginal discharge) was more common in women with a pessary, a RR 1.88 (95% CI 1.57-2.27)[96]. Women with cervical length ≤20 mm also received treatment with vaginal progesterone[96]. Another study could not confirm the reduction of PTD in combination with vaginal progesterone and pessary if the cervical length was ≤25 mm in singleton pregnancies[98]. A Spanish study compared the use of the pessary with progesterone, but could not show any differences in frequency of sPTD <34 GW between the groups, although the prevalence of vaginal discharge was significantly higher in the pessary group (27% vs. 3%, p<.001)[99]. It is still unclear if a pessary can be recommended in singleton pregnancies. A systematic review did not show any reduction of PTD or improved perinatal outcome in singletons[96] (Table 3). Another systematic review could neither show any reduction of PTD <34 GW in twin pregnancies[96] however more recently promising results were published regarding multiple pregnancies where there was a 35% reduction of sPTD<34 GW[97] (Table 3).

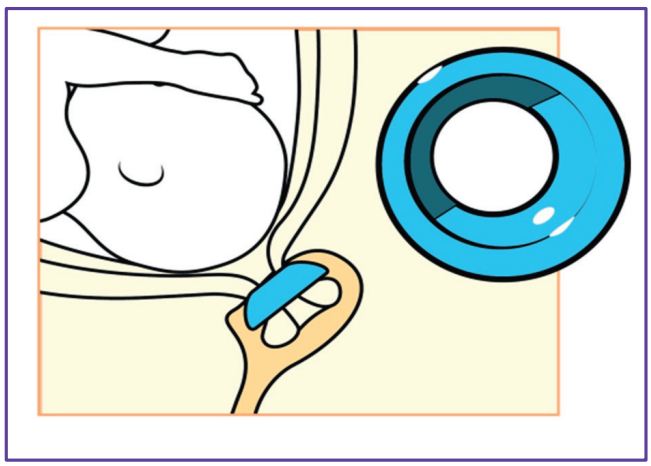


Figure 8 Pessary placed around cervix (permission from @bo.jacobsson)

Progesterone

The effect of progesterone in the prevention of PTD has been studied in multiple RCTs. Intramuscular weekly injections of 17-α- hydroxyprogesterone caproate (250 mg)[100,101] or vaginal natural micronized progesterone (90 mg, 100mg, Crinone®, Lutinus®) have been used[86,87,102-105]. Progesterone influences the myometrium and cervix via progesterone receptors. It supresses smooth muscle contractility and cervical ripening. Trials have included singleton pregnancies with a short cervix and/or risk factors for PTD[86,87,101-105]. Some studies have included twin pregnancies with a short cervix[102,103]. Recently, an individual patient data meta-analysis showed that vaginal progesterone significantly reduced PTD <33 GW by almost 40% (RR 0.62; 95% CI 0.47-0.81, p=0.0006) and composite neonatal morbidity/mortality, RR 0.59 (95% CI 0.38-0.91, p=0.02) in women with a singleton pregnancy and a short cervical length ≤25mm in the second trimester[88]. The most important systematic reviews on progesterone used to prevent PTD are summarized in Table 3. In the OPPTIMUM trial (n=1228) women were randomised to vaginal progesterone or placebo and children follow-up was performed up to two years of age[87]. No significant effect on the child's neurodevelopment between the groups was found.

The administration of progesterone (either intramuscularly or vaginally) for women with twin pregnancies does not appear to be associated with a reduction of PTD or with improved neonatal outcome in a systematic review of 17 trials involving a total of 4773 women[106].

Table 3 Overview of systematic reviews of interventions (cerclage, pessary and progesterone) for prevention of preterm delivery

Systematic reviews Population	Cervical length/ inclusion criteria	Primary outcome	RR (95% CI)	Comment
Cerclage: Systematic reviews (n=2)				
Alfirevic et al **Cochrane Database Syst Rev 2017**[93] Systematic review and meta-analysis 15 studies included, 11 studies including singleton pregnancies, 3 included mixed singleton and multiple pregnancies, 3 compared cerclage with progesterone, others cerclage to expectancy (Althuisius 2001, 2003, Beigi 2003, Berghella 2004, Chandiramani 2010, Èzechi 2004, Ionescu 2012, Keeler 2009, Lazer 1984, MRC/RCOG 1992, Owen 2009, Rush 1984, Rust 2009 Simone 2009, To 2004)		PTD <34 GW (9 RCTs) PTD<37 GW (9 studies, 2898 women) All perinatal loss (10 studies) Serious neonatal morbidity (6 studies) Baby discharges home healthy	**0.77 (0.66- 0.89)** **0.80 (0.69-0.95)** 0.82 (0.65-1.04) 0.80 (0.55-1.18) 1.02 (0.97- 1.06)	Effect of cerclage to reduce PTD <37 and <34, but no effect on neonatal outcomes
Berghella et al UOG 2017[107] Systematic review and meta-analysis Including 5 RCTs. Asymptomatic singleton without prior sPTD n=419 women (Althuisius 2001, Berghella 2004, Otsuki 2016, Rust 2001, To 2004)	≤25mm ≤15mm ≤10mm	PTD<35 GW	0.88 (0.63-1.23) 0.68 (0-57-1.13) **0.68 (0.47-0.98)**	No progesterone treatment No difference in neonatal outcome
Pessary: Systematic reviews (n=2)				
Saccone et al J Ultrasound Med[96] Systematic review and meta-analysis Including 3 RCTs (Goya 2012, Hui 2013, Nicolaides 2016) Total 708 women with pessary and 712 women in control group. Singleton pregnancies, pessary vs expectancy	≤25 mm	sPTD<37 GW sPTD <34 GW sPTD32 GW	0.50 (0.23-1.09) 0.71 (0.21-2.42 1.32 (0.87 – 2.01)	No effect on perinatal morbidity or mortality
Zheng et al J Matern Fetal Neonatal Med 2019[97] Systematic review and meta-analysis Included 11 studies: 4 singletons (3 RCTs, 1 cohort study) (Arabin 2003, Goya 2012, Hui 2013, Nicolaides 2016) and 8 on multiple pregnancies (3 RCTs, 5 cohort), (Arabin 2003, Carreras 2012, Liem 2013, Goya 2016, Nicolaides 2016, Monfrance 2016, Fox 2016, Di Tommaso 2016)	≤15-25 mm	sPTD <34 GW (singleton and multiples) sPTD < 34 GW (singletons) sPTD <34 GW (multiples)	**0.65 (0.44-0.96)** 0.71 (0.21-2.42) **0.64 (0.44-0.96**)	No significant effect on perinatal mortality or composite adverse perinatal outcome in singletons or multiple pregnancies
Progesterone: Systematic reviews (n=2)				
Dodd et al Cochrane database Syst Rev 2013[108] Systematic review and meta-analysis Progesterone for women with short cervix, 1554 women from four studies included (Fonseca 2007: cervix < 15 mm 200 mg vaginal progesterone; Grobman 2012 , cervix <30 mm, 250 mg 17-OH progesterone caproate i.m. 1x/week, Hassan 2011 cervix between 10 and 20 mm, 90 mg vaginal preogesterone; Rozenberg 2012 cervix <25 mm, 500 mg 17-OH progesterone caproate im 2x/week)	≤15mm 10-20 mm <30 mm	**≤34 GW (2 studies)** ≤28 GW (2 studies) <37 GW (3 studies)	**0.64 (0.45- 0.90)** 0.59 (0.37 – 0.93**)** 0.97 (0.82- 1.15)	Major neurodevelopmental handicap was not reported. More uriticaria in progesterone group (one study) No effect on perinatal outcome
Romero et al AJOG 2018[88] Systematic review and meta-analysis of individual patient data (IPD MA), 974 woman, black 36%, white 38%, prior sPTD 29% Five studies included (Fonseca 2007, O´Brien 2007, Cetingoz 2011, Hassan 2011, Norman 2016) Included 974 singleton pregnancies in IPD MA, midtrimester cervical measurement, vaginal progesterone treatment (n=498) vs placebo (n=476)	≤ 25 mm	PTD <33 GW PTD <37 GW sPTD <33 GW sPTD <34 GW Neonatal deaths 1.7% vs 3.2% Adverse perinatal outcome	**0.62 (0.47-0.81)** 0.90 (0.77 – 1.05) **0.70 (0.51- 0.95)** **0.72 (0.55 -0.95)** 0.44 (0.18 -1.07) **0.59 (0.38-0.91**)	No long-term child outcome
CI=confidence interval, GW=gestational weeks, PTD=preterm delivery, RR=relative risk, sPTD=spontaneous prterm delivery				

1.5 Measuring cervical length

Different cervical assessments

Cervical length can be measured with transvaginal, transabdominal or transperineal ultrasound. Comparative studies with transvaginal ultrasound have been performed with transabdominal and transperineal ultrasound[109-117], with digital examination[118,119], and with measurement with a cervicometer[120]. Cervicometer (Cervilenz™) is a stitch which can placed in lateral fornix in speculum examinations. The bulging part of cervix into vagina, ie. distance from lateral fornix to the top of the cervix, is measured with the measuring stitch[120]. Most RCTs are performed with transvaginal ultrasound and it is considered the method of choice in many guidelines to measure the cervical length, however, some prefer transabdominal ultrasound[121-124].

Measurements to evaluate the softness of cervix has been studied with elastography or Hitachi- Real-time Tissue Elastography (HI-RTE)[125-128]. Other examinations performed with transvaginal ultrasound to predict preterm delivery are: cervical consistency index (counted from endocervical longitudinal measures with and without pressure)[129,130], volume of cervix (2D or 3D way)[131-133], uterocervical angle[134,135], tissue specific grey scale analysis[136], evaluation of endocervical glandular area[137,138], vascularization index/vascularization flow-index/flow index measured with 3D transvaginal ultrasound[139,140] or measuring the pulsatile index of arteria uterina with doppler[141].

Measurement technique and criteria for cervical measurement with trasnvaginal ultrasound

The most common definitions of cervical measurement with transvaginal ultrasound are referred to the descriptions according to Iams 1996, Heath 1998 or Kagan 2015[60,64,142]. In some studies, the measurement criteria of cervical length are defined separately[143-145]. Most observational studies have defined the measurement criteria well, in contrast to the RCTs including women with “short cervix”, which have described poorly or not at all how the measurements of cervical length were assessed[86,87,91,103,146]. It makes the most important criteria (a short cervix) for inclusion of women in RCTs a bit uncertain and may be questionable. The measurement criteria, as shown below is a modification from Iams’ and Heath’s criteria with addition of consciousness about lower echogenicity around the endocervical canal and sometimes the presence of the lower uterine segment, e.g. isthmus.

List of the criteria for cervical length measurement with transvaginal ultrasound:

- Lithotomy position
- Emptied urinary bladder
- Sagittal view, cervix occupying 75% of image, 180° angle
- Equal thickness of anterior and posterior lip
- Endocervix visible from external (A) to inner (B) cervical os and correct placement of calipers, low echogenicity in cervical mucosa (Figure 9)
- If isthmic part is visible, it is not included in endocervical measurement (Figure 10)
- Distances are measured as straight line between A-B (endocervix)
- Three repeated measurements, minimum 3 min examination time, still images and video-clips saved, the shortest endocervical measurement is recorded

Figure 9 and 10 shows the placement of the markers[60,64,142,147].

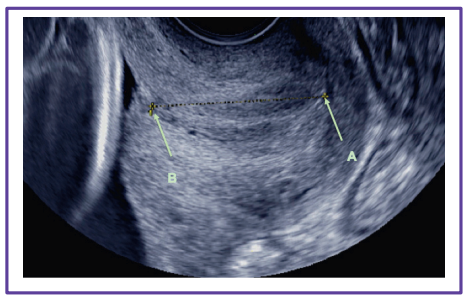


Figure 9 Cervix without isthmus

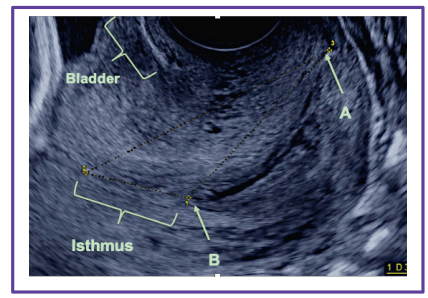


Figure 10 Cervix with isthmus

Learning, training

Measurements from ultrasound examinations are always individual visual assessments based on well-defined criteria including knowledge of anatomy and the performance of ultrasound equipment. Examination technique and a good professional relationship with the patient are also important. The learning process for the medical professional is divided into a theoretical part (with e-learning) and hands-on training[148]. This leads to certification[148]. An online course combined with carrying out a transvaginal cervical evaluation under the supervision of an instructor improved the quality of cervical measurements in residents in obstetrics and gynecology[149]. For certification, assessments are performed according to the certification program and then images are sent for audit. After certification the sonographer can assess the measurements independently. Quality review protocols are proposed for cervical measurements in the same manner as for nuchal translucency measurements[147].

Web based learning and certification programs for transvaginal ultrasound measurement of cervical length are available on the internet (https://clear.perinatalquality.org/, https://fetalmedicine.org/education/cervical-assessment/, https://www.isuog.org/resource/how-to-measure-cervical-length.html). One study showed a positive effect on image quality in those who had followed an e-learning module on cervical measurement, compared with those who had not[148]. YouTube provides a lecture on cervical measurement by MD Celeste Sheppard (https://www.youtube.com/watch?v=SbMKkQ7QaOs).

Pitfalls

Two certification programs (Cervical Length Education and Review program, CLEAR, and Maternal-Fetal Medicine Units, MFMU) have reported certification rates of 85% and 83% respectively at the first attempts[147,150]. The two most common criteria missed by those who did not pass the certification exam were the inadequate enlarging of the image i.e cervix does not fill 75% of the image[147], and too much pressure on the anterior lip leading to unequal thickness of the lips[147,150]. Other difficulties were incorrect caliper placement, the urinary bladder not being empty or insufficient examination time.

A long and curved cervix, as well as cysts or myomas in the cervix, can create shadows and so limit image quality and make the assessment of cervical length difficult[151]. A low-lying placenta near or over the internal os, or the presence of isthmus, may complicate the identification of the internal os[142]. Identification of the outer os is usually easier but it can sometimes be difficult to distinguish it from the posterior vaginal wall.

When cervical measurement is used to decide to treat or not, a dichotomous cut-off value is often used, for example at ≤ 25 mm. One study of the distribution of cervical measurements noted a dip at the cut-off value[152]. They suggested that a predefined cut-off value for a short cervix influences the distribution of the cervical length measurements and may result in fewer measurements around the cut-off value[152].

Cervical measurement using transvaginal ultrasound in the second trimester in asymptomatic women for prediction of preterm delivery – a systematic review of the literature

A systematic literature search was carried out on PubMed on 2018-11-21 to find all observational studies on cervical measurements and their ability to predict PTD, using transvaginal ultrasound in the second trimester. The following search terms were used:

((((("cervical length measurement"[MeSH Terms] OR "cervical length"[tiab] OR "Cervix length"[tiab] OR SCL[tiab] OR "short cervix"[tiab] OR "short cervical length"[tiab] OR "uterine cervical incompetence"[mh] OR "cervix"[tiab] OR "cervical"[tiab]) AND ("premature birth"[MeSH Terms] OR "pre-term"[tiab] OR "obstetric labour"[All Fields] OR "preterm"[tiab] OR "spontaneous preterm birth"[tiab] OR "premature"[tiab])) AND ("sonography"[tiab] OR "ultrasound"[tiab] OR "ultrasonography"[mh])) AND ("transvaginal"[tiab] OR "Cervix Uteri/diagnostic imaging"[MeSH Terms] OR (("diagnostic imaging"[Subheading] OR ("diagnostic"[All Fields] AND "imaging"[All Fields]) OR "diagnostic imaging"[All Fields] OR "ultrasonography"[All Fields] OR "ultrasonography"[MeSH Terms]) AND Prenatal[All Fields] AND "methods"[MeSH Major Topic]))) AND ("prediction"[tiab] OR "risk"[tiab] OR "diagnosis"[tiab] OR "screening"[tiab] OR "sensitivity"[tiab] OR "scan"[tiab] OR "significance"[tiab])) AND (hasabstract[text] AND ("1990/01/01"[PDAT] : "3000/12/31"[PDAT]) AND "humans"[MeSH Terms] AND (English[lang] OR Finnish[lang] OR Norwegian[lang] OR Swedish[lang] OR Danish[lang]) AND "female"[MeSH Terms])) NOT ("twins"[All Fields] OR "pregnancy, multiple"[MeSH Terms]) AND (hasabstract[text] AND ("1990/01/01"[PDAT] : "3000/12/31"[PDAT]) AND "humans"[MeSH Terms] AND (English[lang] OR Finnish[lang] OR Norwegian[lang] OR Swedish[lang] OR Danish[lang]) AND "female"[MeSH Terms])

Inclusion criteria were: 1) prospective cohort studies or systematic reviews, 2) cervical length evaluated with transvaginal ultrasound in asymptomatic women with singleton pregnancy in the second trimester between 14 and 24 GW, 3) participating women considered at low risk of spontaneous preterm delivery or from the general population, and 4) English and Scandinavian languages were accepted.

Results of systematic literature search

The search retrieved 520 articles. After abstract screening 59 full text articles were identified. Twenty-one articles of these fulfilled our inclusion citeria and were included: 3 systematic reviews[71,153,154] and 18 original articles. Original articles are divided into blind or not blind. A study was defined as blind if participants and caregivers were blind to the results of the cervical length measurement. Of the included studies, six were blind, three were unclear whether they were blind or not, and nine were not blind.

A summary of the blind[64,68,69,76,155,156] and not blind studies[55,59,66,67,77,138,157-162] are presented below and in Table 4 and 5

Summary of blind studies

None of the blind studies compared the study population with the general population and selection bias can therefore not be excluded[64,68,69,76,155,156]. Two studies excluded women with indicated PTD in the analysis[69,155], while three studies excluded women with risk factors for PTD[64,155,156], including conception after assisted reproductive technology, prior conization or prior PTD.

A cervical length ≤25 mm had a sensitivity to predict sPTD <35 GW of between 7% and 39.1%, and a specificity of between 92.2% and 100%. PPV varied between 12.9% and 15%. A cut-off of cervical length ≤25 mm corresponded to between 0.3th and 10th percentile. Diagnostic accuracy is shown in Table 4.

Table 4 Summary of blind observational studies

Study	Population	Gestational week	Cervical length	Sensi-tivity	Speci-ficity	PPV	LR+	LR-	AUC
Carvalho et al. *UOG* 2003[76]	Single center, Sao Paolo, Brazil 529 women (34.6 % white, 18. 7% smoker, 7.6% had prior sPTD)	sPTD <33 GW sPTD <35 GW	≤20 mm	40	97	23.5	13.1	0.62	
				42.3	96.7	37.9	7.4	0.60	
Davies et al. *JOGC* 2008[155]	Single center, Kingstone, Canada 963 women (9.2% had prior sPTD)	PTD <37 GW 4.8% (46)	≤20 mm 0.5% (5)	2.2	99.8	20	5.0	0.98	
			≤25 mm 3.2% (31)	13	97.3	12.4	4.8	0.89	
			≤30 mm 14.7% (142)	26.1	85.8	8.5	1.8	0.86	
		PTD <35 GW 1.7% (16)	≤20 mm 0.5% (5)	6.3	99.6	20.0	14.8	0.94	
			≤25 mm 3.2% (30)	25	97.1	12.9	8.8	0.77	
			≤30 mm 14.7% (142)	50	85.9	5.6	3.5	0.58	
Iams et al. *NEJM* 1996[64]	Multicenter, 10 sites, USA 2915 women (16% had prior PTD)	sPTD <35 GW 4.3% (126)	≤20 mm (22 (5th))	23	97	25.7	7.7	0.8	
			≤25 mm 10th (26)	37.3	92.2	17.8	4.8	0.7	
			≤30 mm 25th	54	76.3	9.3	2.1	0.6	
Leung et al. *UOG* 2005[68]	Single center, Hong kong, China 2880 women, general population (6.9% prior PTD, 100% chinese)	PTD <37 GW							0.56
		PTD <34 GW	≤20 mm 0.2nd	10.5	99.9	40.0	100.4	0.9	0.68
			≤25 mm 1.8th	26.3	98.3	9.4	15.7	0.7	
			≤30 mm 10th	36.8	90.1	2.4	1.8	0.7	
			≤35 mm 35th	63.2	65.5	1.2	1.8	0.6	
Matijevic et al. *Int J Gynaecol Obstet* 2010[156]	Single center, Zagreb, Croatia 316 low risk women	sPTD <34 GW 2.5% (8)	≤26 mm 5th	87.5	97.4	46.6	33.7	0.13	
		sPTD <37 GW 7.2% (23)		47.8	98.6	73.3	46.7	0.53	
Taipale and Hiilesmaa. *Obstet Gynecol* 1998[69]	Single center, Espoo, Finland, one examinator, one controller 3694 women (99% white), (n=20)	sPTD <35 GW 0.8% (31)	≤25 mm 0,3rd	7	100	15	-	0.9	
			≤29 mm 3rd	19	97	6	6.3	0.8	
			≤35 mm 27th	45	73	1.4	1.7	0.8	
		sPTD <37 GW 2.4% (88)	≤25 mm 0,3rd	6	100	39	-	0.94	
			≤29 mm 3rd	16	97	13	5.3	0.9	
			≤35 mm 27th	35	73	3	1.3	0.9	

AUC= area under curve (Receiver operation characteristics curve), GW= gestational week, LR-=negative likelihood ratio, LR+= positive likelihood ratio, NPV= negative predictive value, PPV= positive predictive value, PTD= preterm delivery, sPTD= spontaneous PTD

Summary of non blind studies and studies with unclear blinding

Studies were not blind if they recruited women to RCTs: Hibbard[161], (cerclage if CL ≤25mm), Kazemir[77] (Dutch population, progesterone, if CL ≤30), van der Ven[65] (progesterone, if CL≤ 30 mm). In some studies women with cervical length below a certain cut-off point were treated with progesterone, cerclage or pessary decided by the clinician[55,66,67,77,157,158]. The study from Malaysia was not blind if cervical length was <20 mm because the ethical committee did not approve it[160].

Wulff et al[55] compared the study population with women who had declined to participate in the study and revealed more risk factors in the study population than among decliners. Van der Ven[65] compared maternal background data of the study population with both the estimates from the Dutch population as a whole and with those in the study population who had missing delivery data. Both of these populations had more nulliparous women than the study population. Delivery outcome data from the study population were compared with national delivery outcome data in two studies[66,161], and both reported fewer women with PTD and sPTD in the study groups. Only four studies report the number of women who declined to participate[55,67,138,159].

A cervical length ≤25 mm could predict sPTD <37 GW with a sensitivity between 8% and 77%, a specificity of between 95% and 99%, and PPV varied between 7% and 56%. AUC was between 0.51 and 0.91. Diagnostic accuracy is shown in Table 5.

Table 5 Summary of non-blind studies and studies with unclear blinding

Author, journal and year	Population, country, ethnicity, IVF, prior PTD	Gestational week (PTD/sPTD) Prevalence % (n)	Cervical length cut-off mm percentile/ prevalence % (n)	Sensi-tivity	Speci-ficity	PPV	LR+	LR-	AUC
Barber et al. ***Int J Gynaecol Obstet*** **2010**[159]	Single center, Las Palmas de Gran Canaria, Spain 2351 women	PTD<37 GW 7.2% (184)	≤28mm 3rd	26	98	63.6	13	0.35	-
			≤29mm 5th	34	97	51	11.3	0.66	-
			≤30mm 10th	39	91	31	5.6	0.67	-
Dilek et al. ***Gynecol obstet invest*** **2007**[162]	Single center, Mersin, Turkey, single examinator 257 women, low risk	PTD <37 GW 7.4% (19)	35.3 mm (at 16 GW)	26.3	93.7	25	11.4	0.79	0.57
			34.3 mm (at 24 GW)	84.2	81.5	26.7	4.6	0.19	0.91
Esplin et al. ***Jama*** **2017**[66]	Multicenter, 8 centers USA 8871 women at visit 2. (60.7% white, 5.0% had prior sPTD)	sPTD <32 0.76% (67)	≤20mm	14.9	98.8	8.6	12.3	0.86	0.57
			≤25mm	23.9	97.9	7.4	10.4	0.78	0.61
		sPTD <37 GW 5.0% (439)	≤20 mm	4.1	98.8	15.5	3.5	0.97	0.51
			≤25 mm	8.0	98.8	16.2	3.67	0.94	0.53
Heath et al. ***UOG*** **1998**[67]	Single center, London, UK 1232 with cervix >15mm and 21 with cervix ≤15mm and no cerclage (expectance), (white 47.6%, smokers 14.7%, 3.8 % prior sPTD, nullipara 30.4%)	sPTD <27 GW 0.3%	≤15 mm 1.7% (43)	100					
		sPTD <33 GW 1.5%		58					
		sPTD <37 GW 5.0%		20					
		sPTD <33 GW 1.5% (19)	≤20 mm 3.4%	57.9	92.9	2.8	8.15	0.45	-
			≤30 mm 18%	68.4	62.9	11.2	1.29	0.50	-
Heath et al. ***BJOG*** **2000**[157]	Single center, London, UK 5068 women (white 50%, smokers 14.7%)	sPTD <33 GW 0.8% (42/5069)	≤15 mm 1.5%	27.9	99.5	30.8	55.8	0.72	-
Hebbar and Samjhana. ***Medical J Malaysia*** **2006**[160]	Single center, Manipal. Malaysia 168 women	sPTD <37 GW 7.7% (13/168)	≤25 mm	77	95	56	15.4	0.24	
Hibbardet al. ***ObGyn*** **2000**[161]	Single center, Chicago, USA 760 women (77.2%black, 14.8% white)	sPTD <32 GW 3.5% (27/760)	<22 mm 2.5th	18.5	97.9	27.0	8.8	0.11	-
			<27 mm 5th	29.6	95.8	22.9	7.0	0.14	-
			<30 mm 10th	44.4	89.9	14.8	4	0.23	-
		sPTD <35 GW 6.7% (51)	<22 mm 2.5th	21.6	97.7	47	9.4	0.11	-
			<27 mm 5th	29.4	96.4	43.9	8.2	0.12	-
			<30mm 10th	41.2	90.7	27	5.0	0.23	-
		sPTD <37 GW 11.2% (85)	<22 mm 2.5th	12.9	98.5	30	8.4	0.12	-
			<27 mm 5th	20	96.6	46.9	5.9	0.17	-
			<30 mm 10th	32.9	91.3	32.7	3.8	0.26	-
Jwala et al. ***AOGS*** **2016**[158]	Single center, Seattle, USA 528 women (32% white, 55% african-american, 51% nulliparous)	sPTD <37 GW 6.82% (36)	≤20 mm (6)	11.1	99.6	66.7	27.8	0.04	-

Kazemier et al. *J Perinat* 2016[77]	Two cohorts 3409 from The Netherlands White 86.9%, smoking 0.6%, IVF 3%, prior PTD 8%	sPTD <32 GW 60 (1.8%)	≤25 mm 0.9%						0.65
		sPTD <37 GW 198 (6.3%)							0.63
	3334 from Chicago, USA White 61.3%, smoking 0.7%, IVF 5.1%, prior PTD 7.3% From rutine transvaginal ultrasound screening, not blinded, treatment offered	sPTD <32 GW 33 (1.0%)	≤25 mm 0.8%						0.72
		sPTD <37 GW 175 (5.3%)							0.64
Pires et al. *Int J Gynaecol Obstet* 2006[138]	Single center, Sao Paolo, Brazil 338 women	sPTD <35 GW 3.2% (11)	<20mm	27.3	97.9	30	13	0.74	-
		sPTD <37 GW 6.2% (21)	≤20mm	18	98	40	9	0.63	-
van der Ven et al. AOGS 2015[65]	Multicenter study, 59 centers, The Netherlands. 11943 women divided to 5710 nulliparous and 6233 (low-risk multiparous, 88% white)	sPTD <37 GW 464 (3.9%) PTD <37 GW 666 (5.6%)							
	5710 Nulliparous Included: 30 (0.5%) women with progesterone.	sPTD <32 GW 0.70% (40)	≤30	12.5*					
		sPTD <34 GW 1.3% (77)	≤30	10.8	98		5.4*	0.91*	0.63
			<35	33.1	86.6				
		sPTD <37 GW 5.3% (300)	≤30	6.3*	96.1*	15.2*	1.62*	0.98*	0.61
			≤35	28.2	87.3				
	6233 Low-risk multiparous Included: 11 (0,2%) with Progesterone Excluded: multiparous with prior sPTD<34	sPTD <32 GW	≤30	13.3*					
		sPTD <34 GW 0.4% (25)	≤30	9.1	96.7				0.58
			<35	23.6	90.2				
		sPTD <37 GW 2.6% (164)	≤30	5.4*	98.7*	10.6*	4.23*	0.96*	0.56
			<35	10.0	90.4				
Wulff et al. *UOG* 2018[55]	Multicenter, Denmark 3442 women (97.5% white)	sPTD <28 GW 0.1% (4)	≤25 mm 0.78% (26)	75*	99*	11.5*	112*	0.25	
		sPTD <34 GW 0.8% (29)	≤25 mm 0.78% (26) ≤25 mm 1.79% (59)	20.7*	99.4*	23*	34.2*	0.80*	
		sPTD<37 GW 3.3% (114)		8.8*	99.5*	38.5*	17.7	0.92	
*counted from the data in the article									

1.6 Intra-observer and inter-observer reproducibility and reliability

General methods to study reproducibility and reliability

The reproducibility of a measurement can be studied by two or more observers, which is called inter-observer reproducibility, or by the same observer, which is called intra-observer reproducibility. Depending on whether the measurement result

is a categorical (e.g. ≤25 mm cut-off of cervical length, or isthmus or not in our study) or a continuous variable, different statistical methods are used. Observers should be blinded from each other's (inter-observer) or from their own (intraobserver) measurements, and the equipment and the scale used should be the same.

A typical way to study differences between observers is to see whether the absolute inter-observer differences increase with increasing measurement value or not by calculating Spearmans rank correlations coefficient. If no significant correlation is observed, Figure 11a, the differences are plotted against the mean difference. The Limit of Agreements (LoA) (mean difference ± 1.96 SD) shows in dotted lines in Figure 11b, a Bland-Altman plot. LoA is an interval between mean difference where 95% of future measurements by two examiners, working as a pair, are expected to fall.

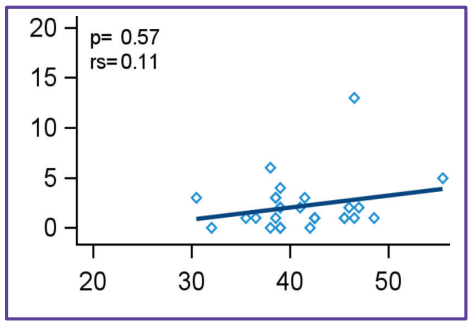


Figure 11a

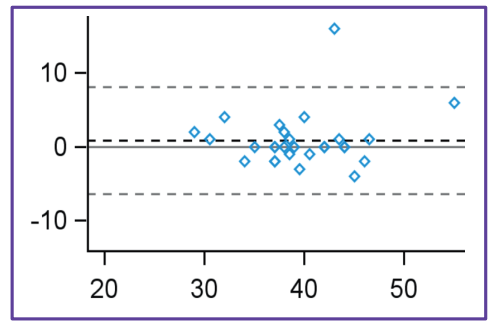


Figure 11b

Figure 11a Inter-rater Bland-Altman plot with Spearman correlation coefficient Absolute difference vs mean difference, for all pairs of examiners (from LIVE study, Paper III).
Figure 11b Inter-rater Bland-Altman plot: Absolute differences against mean difference and limits for agreement (LoA), total for all raters (From LIVE study, Paper III). Dotted lines are mean and LoA.

An overview of statistical methods used in reproducibility studies are listed in Table 6.

The common terms used in reproducibility studies are[163]:

- Agreement means the degree to which scores or ratings are identical among the observers/examiners
- Repeatability is a degree of how close scores or ratings obtained under similar conditions are.
- Reliability may be explained as a ratio of variability between subjects (e.g. women) or objects (e.g. transvaginal ultrasound) to the total variabilty of all measurements in the sample. And, therefore, can be defined as the ability of a measurements to differentiate among subjects or objects.

Table 6 An overview of the most important statistical methods in reproducibility studies

Method	Abbreviation	What it measured and how?	Interpretation
Inter-observer measurement, two (or more) observers assess the measurement of the same subject under similar conditions			
Bland-Altman plot		Mean difference between observers (in y-axis) against the mean measurement value (x-axis).	Graphical way to see if differences between observers increase in increasing measurement values and how the differences are distributed around the mean
Mean difference		Observer 1 measurement - observer 2 measurement = difference between observers Mean difference = sum of differences / number of observations	The difference between observations (in mm etc) in a pair
Limits of Agreement	LoA	LoA= mean difference +/- 1.96 x $SD_{diffenence}$ 95% of differences between the two examiners fall between these limits in future measurements	Prediction interval for observer 2 given a measurent for observer 1
Intra-observer measurement, the same observer repeats the measurement again under similar conditions			
Intra-individual standard deviation= measurement error	IISD	$\text{IISD} = \sqrt{\frac{\Sigma d_i^2}{2n}}$ d_i = difference between two observations	Measurement error, difference between a subject's measurement and the true value is expected to be less than 1.96 xIISD for 95% of observations[164]
Repeatability		$\sqrt{2} \times 1.96 \times \text{IISD} \approx 2.77 \times \text{IISD}$	The difference between two measurements of the same subject is expected to be less than IISD x 2.77 in 95% of pairs of observations[164]
Inter and intra-observer reliability			
Intra-Class Correlation (ANOVA two-way random model)	ICC	Express inter and intra-observer reliability The ICC is the proportion of variance between examined individuals and the total variance Used for continous variables	If ICC is near 1, it has high reliability If ICC is near 0, it has low reliability
Cohen's kappa		Express agreement adjusted for that expected by chance Used for dichotomous variables	The range of Cohen's kappa is between 0 and 1, values near 1 are excellent

Reproducibility studies of cervical measurement using transvaginal ultrasound in the second trimester – a systematic review of the literature

To find all reproducibility studies performed on cervical measurements using transvaginal ultrasound in the second trimester, a systematic literature search was carried out in PubMed, Embase and Cochrane library in 2018-12-21. No language or time restrictions were applied. Search strategy, flow chart and a list of excluded articles with reasons for exclusion are shown in Appendix pp 3-7 in Paper III.

Only two papers were included which presented data on reproducibility in cervical length measurement using transvaginal ultrasound in the second trimester, França et al[165] and Heath et al[60].

França et al studied inter- and intra-observer reproducibility in the first and second trimester. In the second trimester, 31 women were examined by three of four doctors one after each other and each doctor made three measurements in a live situation. Inter-observer reproducibility was studied in pairs (Observer 1-2, 1-3, 2-3, 2-4, 1-4, 3-4). Intra-observer reproducibility was evaluated by examining each observer's three measurements. Intra-observer reliability was high, ICC being 0.90 (between 0.85 and 0.91). The inter-observer ICC was 0.58 (min 0.53, max 0.64)[165]. The width of the LoA ranged from 12 mm to 23 mm, which is similar to a study by Valentin and Bergelin, who had 21 mm for one pair when measuring cervical length in the second and third trimester[166].

In the study by Heath et al the cervical length of 100 women was measured by two of four observers (well-trained operators). The first observer made two live measurements, then the other observer made a measurement of the first observer's second frozen image, where the calipers had been removed. The second observer then made two of his/her own live measurements.The first observer then made a measurement of the second observer's second image. Thus six measurements were obatined on each woman. The number of examiner pairs is not mentioned. The observers were blind to all results. The standard deviation of the intra-observer live measurements was 1.76 mm, and of the inter-observer measurements 2.13 mm. On 95% of occasions the difference between two measurements of the same observer (intra-observer) was ≤3.5 mm and between two observers (inter-observer) ≤4.2 mm[60].

Both studies by Heath and França[60,165] included few examiners (n=4), and so the generalizability of the results may be questioned. The profession of the people performing the measurements was unclear in the Heath study. When the accuracy of measurements is based on only a few observations for each pair, the general conclusions cannot be applied to clinical practice or compared with multicenter studies where there can be 30 to 100 different examiners from different professions. It is desirable that the LoA should not increase ±5 mm, but in a clinical context it may be necessary to accept higher LoA because of the heterogenicity of both the examiners and patients.

2. Aims

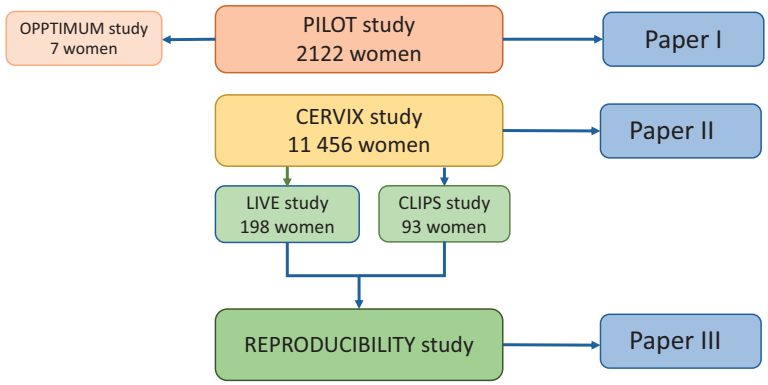
The overall aim of this thesis is to measure the cervical length of asymptomatic women in their second trimester in Sweden, using transvaginal ultrasound. It is also to investigate the relationship between cervical length and PTD and to study the agreement and reliability of the measurement method in the hands of several examiners.

The specific aims are to evaluate:

- cervical length as a risk factor for PTD (Paper I and II)
- the prevalence of a "short cervix" in Sweden (Paper I and II)
- the optimal cut-off of cervical length to predict PTD (Paper I and II)
- the optimal time period to measure cervical length in order to predict PTD (Paper II)
- whether the difference between two measurements is more predictive than one measurement of cervical length (Paper II)
- the discriminative ability of "short cervix" to predict sPTD <34 GW (Paper I) and PTD <33 GW (Paper II)
- the discriminative ability of "short cervix" to predict sPTD <28, <29, <30, <31, <32, <33, <34, <35, <36, and <37 GW (Paper II)
- agreement repeatbility, and reliability of cervical measurements, when using transvaginal ultrasound (Paper IlI)
- agreement regarding the presence /absence of the isthmic part of the cervix uteri and regarding the cervical length ≤25mm or >25 mm (Paper IlI)
- whether cervical screening is a method to apply in Sweden and if an RCT of prophylactic treatment with progesterone is realistic (Paper II)

3. Patients and methods

Overview of Paper I-III



3.1 Study populations

Paper I

Women attending the routine ultrasound scan in the their second trimester were recruited for a study to measure cervical length using transvaginal ultrasound at Sahlgrenska University Hospital, Gothenburg and Norra Älvsborg County Hospital, Trollhättan in Sweden. Certified midwives measured endocervical length between 16+0 GW and 24+0 GW. Three measurents of endocervical lengths were performed and the shortest length was recorded in the medical record. The study lasted between August 2012 and May 2013. Women with fetal malformations, multiple pregnancies, age <18 years, or signs of ongoing miscarriage (bleeding, leakage of amniotic fluid) were not included. Women with a cervical length ≤25 mm were recruited to participate in an RCT and were randomised to vaginal progesterone or placebo as treatment for prevention of PTD (OPPTIMUM study[87]). Women with a cervical length >25 mm followed the ordinary antenatal care program. Women with singleton deliveries at Sahlgrenska University Hospital during the study period, who were not screened for cervical length, consituted a control group. The sample size was calculated in order to recruit 750 women with cervical length ≤25 mm to a Swedish

RCT on vaginal progesterone. The prevalence of cervical length ≤25 mm was estimated to be 2.5% so cervical length measurements of 10,000 women were needed annually over 3.5 years. After ten months only seven women were recruited and the RCT was discontinued.

Paper II

Women attending the routine ultrasound scan in the second trimester were recruited for a study to measure cervical length with transvaginal ultrasound at seven centers in Sweden (Sahlgrenska University Hospital [Gothenburg], Skåne University Hospital in Malmö and Lund, Solna and Huddinge Hospitals, Karolinska Institute [Stockholm], Falu Hospital [Falun] and Örebro University Hospital) between May 2015 and June 2017. This was called the CERVIX study. Women with fetal malformations, multiple pregnancy, age <18 years, difficulties to understand oral and written information, gestational age <18+0 GW or >20+6 GW, or earlier participation in the study were not eligible. Women with signs of ongoing miscarriage (bleeding, leakage of amniotic fluid, amniotic sac bulging into vagina), using progesterone or cerclage were excluded. Cervical length was measured twice: the first between 18+0 and 20+6 GW (Cx1) and the second between 21+0 and 23+6 GW (Cx2) with a minimum of 14 days between the measurements. Both the women and the health care providers were blind to the results of the cervical length measurements. Women who declined cervical length measurement but consented to data collection comprised one of the two control groups, the No Cervix measurement population. The other control group consisted of a Swedish background population from the Pregnancy Register. The Background population comprised all women ≥18 years old with a singleton pregnancy delivering in Sweden during the study period (the first singleton delivery occurring during the study period). Sample size was calculated with the aim of getting at least 100 women with PTD <33 GW, which could give a reasonable confidence interval to assess the sensitivity of cervical length in predicting PTD. The study population then needed to include 11,000 women.

Paper III

Paper II consists of two parts, the LIVE study and the CLIPS study. The main purpose of the LIVE study was to estimate the inter-observer agreement and reliability of cervical length measurement. In the LIVE study, thirty women participating in the CERVIX study in each seven centers were asked to have their cervical length measured by transvaginal ultrasound by two midwives, one after the other. In the LIVE study in each examiner pair, the first examiner's measurement of cervical length was included in the CERVIX study. In the CLIPS study, the main purpose was to estimate intra-observer agreement (measurement error and repeatability) and reliability. Sixteen midwives from the CERVIX study individually measured cervical length twice at least two months apart but in a different order on 100 video clips from participants in the CERVIX study. The examiners were blinded to each other's measurements and in the CLIPS study also to their own previous results. The LIVE and CLIPS studies were performed between January 2016 and June 2017.

Sample size was not calculated in the reproducibility studies but was based on the availability of examiners in the participating centers.

3.2 Transvaginal ultrasound measurement of cervical length

Criteria for cervical length measurement (Figure 12)

We measured cervical length in the same manner as described in Chapter 1.4 p 37 and:

- we measured the endocervical length from A to B in Papers I-III and added the following in Papers II and III
- we measured isthmus length from B to C. The point C is called the "virtual inner os" and is the innermost end of the juxtaposed anterior and posterior isthmus
- we measured the straight distance between A and C
- we calculated the sum of A to B and B to C, i.e. endocervical length and isthmus length
- if isthmus length >0 mm in at least one of three measurements, it was denoted as isthmus being present

- if cervical length was not measurable, the reason was given

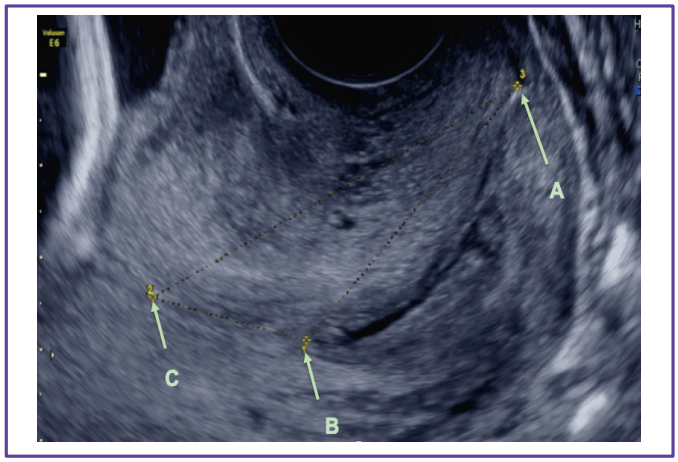


Figure 12 Measurement of cervical length when isthmus is present. A denotes the external os, B denotes the internal os. C (called the "virtual inner os") is the innermost end of the juxtaposed anterior and posterior isthmus. Measurements are taken as a straight line from A to B (the endocervical canal), B to C (isthmus length), and A to C

Certification and quality control

All cervical length measurements were performed by trained and certified midwives. They had two theoretical lectures; one from the Fetal Medicine Foundation (www.fetalmedicine.org/cervical-assessment-1) and one from the steering committee of the CERVIX study. Practical training consisted of hands-on training supervised by a physician or by a certified midwife. For certification the midwife needed to produce ultrasound images fulfilling the quality criteria as shown below from five consecutive women examined at 18 to 23 GW. The midwife was certified if all five members of the quality control committee of the CERVIX study agreed that the images fulfilled the quality criteria.

- Quality criteria for correct transvaginal cervical length measurement
- Cervix occupies at least 75% of the screen
- Anterior and posterior lip of the cervix are of equal thickness
- Full endocervical canal is clearly seen
- Inner and outer cervical os are clearly seen as well as the virtual inner os if isthmus is present Calipers are correctly positioned

In the CERVIX study, regular quality controls were carried out twice a year without prenotification. Images from five consecutive women examined during a specified week were selected retrospectively by each midwife and submitted to the quality control committee. Also pre-planned controls were performed twice a year. The midwives selected three different ultrasound images (from three different women) considered by themselves to meet all the quality criteria. The quality control committee asked for new images if they did not meet the criteria. After three failed attempts the midwife was no longer allowed to examine women in the CERVIX study.

3.3 National registers

For Paper II data was collected on the study populations from:
The Swedish Pregnancy Register
Information on pregnancy, delivery and neonatal data of the index pregnancy
The Swedish Medical Birth Register
Information on previous preterm deliveries before the index pregnancy
The Swedish National Patient Register
Information on previous cervical conization and cerclage during index pregnancy
The Swedish Prescribed Drug Register
Information of redeemed prescription of vaginal progesterone during the index pregnancy

The Pregnancy Register is a national quality register. The Swedish Medical Birth Register, the National Patient Register and the Prescribed Drug Register are health data registers hosted by the Swedish National Board of Health and Welfare (www.socialstyrelsen.se). Linkages between the registers are possible via the unique personal identity number given to all citizens in Sweden.

The Pregnancy Register
Since 2013, data on pregnancy, ultrasound examinations during pregnancy, delivery and neonates has been collected in the Pregnancy Register[167]. The Pregnancy Register covered 93% of all deliveries in Sweden in 2017. Data is obtained from three different sources: 1. Manually entered data by antenatal care midwives at registration for antenatal care (variables which are not registered in electronic medical records) 2. Data on first trimester combined ultrasound and biochemistry 3. Electronic transfer of data from the electronic medical records from 2013 and

onwards. The electronic data transfer includes data from antenatal care, second trimester ultrasound examinations, delivery and postpartum care for the mother and infant, including diagnoses and procedures. The Pregnancy Register was started by merging the Maternal Health Care Register (established 1999) and the National Quality Register for Perinatal Diagnosis (established 2004). The Maternal Health Care Register database has been validated and has very good validity (>95%) in more than 50% of variables and good validity (70-90%) in the remaining variables[168]. Individual, de-identified and aggregated data from the Pregnancy Register is available for care providers and researchers.
The Pregnancy Register publishes annual reports about maternity care and deliveries, the latest report being published in 2018 for the year 2017 (https://www.medscinet.com/gr/uploads/hemsida/dokumentarkiv/GR_Årsrrapport_2017_4.0.pdf).

The Medical Birth Register
The MBR has collected data on pregnancies, deliveries and neonates since 1973 and includes about 99% of deliveries in Sweden[169]. The data was validated in 1988 and 2001 and is considered to have good validity[169], www.socialstyrelsen.se/publikationer2002/2002-112-4). Maternal diagnoses were correctly reported in 98% of cases, data on onset of the delivery was missing in 3.4% (spontaneous onset 82.9%, induction of labour 8.3% and cesarean section 5.3%), and some had more than one onset marked for deliveries between 1992 and 1998. Individual de-identified and aggregated data is available for researchers with a delay of about one year. The pregnancy, delivery and neonatal data from 2016 was published in January 2018 and data from 2017 is not yet available (2019-03-15).

The National Patient Register
The National Patient Register started in 1987 and has collected data on in and outpatient diagnoses according to the International Statistical Classification of Disease and Related Health Problems (ICD)-9 between 1987-1996 and ICD-10 from 1997. Primary health care is not yet included. Validity was reported in 2011 and data was shown to have high validity for most of the diagnoses[170].

The Prescribed Drug Register
The Prescribed Drug Register has collected individual-based data from redeemed prescriptions since 2005. It contains information on Anatomic Therapeutic Chemical (ATC) -code, prescription amount, date of prescription and date when the product was redeemed (www.socialstyrelsen.se). All data is collected electronically and there is little missing data. Drugs used during hospitalization are not recorded.

3.4 Overview of methods in Papers I-III

Table 7 Overview of the studies and statistical methods in Papers I-III

	Paper I	Paper II	Paper III
Study design	Observational, prospective study, two centers	Observational, prospective multicenter study, seven centers	Reproducibility study, prospective, multicenter study, seven centers
Sample size	Study population 2122 women Control group 7216 women	Study population 11 486 women Control groups 9799 women 347 398 women	LIVE: 7 examiner pairs, 24-30 women each CLIPS: 16 raters, video clips from 93 women
Study period	2013 - 2014	2014 - 2017	2016–2017
Data sources and cross-linking	eCRF, Obstetrix*	eCRF, PR, MBR, NPR, PDR	eCRF, Obstetrix*
Statistical methods			
Descriptive continuous variables	Mean, SD, median, minimum, maximum	Mean, SD, median, minimum, maximum, interquartile range	Mean, SD, median, minimum, maximum, interquartile range
Descriptive categorical variables	Numbers, percentage	Numbers, percentage	Numbers, percentage
Analyses of ordered categorical variables	Mantel-Haenszel chi-squared test		
Analyses of continuous- variables	t-test		
Association between dependent and independent variables	Univariable and multivariable logistic regression analysis, OR with 95% CI, adjusted OR with 95% CI, RR with 95% CI	Univariable logistic regression analysis (results not showen)	
Diagnostic test	Sensitivity, specificity, PPV, NPV, LR+, LR-, ROC, AUC	Sensitivity, PPV, ROC, AUC, NNH, NNS	
Survival analysis		Kaplan-Meier plots	
Test of correlation			Spearman rank correlation coefficient
Agreement of continuous variables			Bland-Altman plots. Mean difference and Limits of Agreement (mean difference ±1.96 SD), Width of Limits of Agreement
Agreement of dicotomous variables			Percent total agreement, percent positive and percent negative agreement. Cohen's kappa
Inter and intra-observer reliability of continuous variables			Intra-class correlation coefficient (ICC), ANOVA two-way random model
Inter and intra-observer reliability of dicotomous variables			Cohen's kappa, Fleiss kappa
Intra-observer measurement error			Intra-individual standard deviation (IISD). The difference between a

			subject's measurement and the true value is expected to be less than 1.96x IISD for 95% of observations
Intra-observer repeatability			The difference between two measurements on the same subject is expected to be less than 2.77xIISD (√2x1.96xIISD) for 95% of observations

AUC, area under the receiver operating characteristics curve; CI, confidence interval; eCRF, electronic Case Record Form (MedSciNet AB, Stockholm, Sweden, www.medscinet.com); ICC, intraclass correlation coefficient; IISD, intra-individual standard deviation; LR, likelihood ratio; MBR, The Swedish Medical Birth Register; NNH, number needed to harm, i.e. number of false positive results per one true positive test result; NNS, number needed to screen, i.e. number of women needed to screen to detect one true positive test result; NPR, the Swedish National Patient Register; NPV negative predictive value; PDR, The Swedish Prescribed Drug Register; PPV positive predicitive value; PR, The Swedish Pregnancy Register; ROC receiver operating characteristics curve; SD, standard deviation
*Obstetrix: an electronic medical record system (Cerner AB Sweden)

3.5 Ethical permissions and considerations

The studies were approved by the Regional Ethical Review Board in Gothenburg for Paper I: Dnr 311-12 and for Papers II and III: Dnr 825-13, T053-14, T691-14, T972-15, T122-16, T896-17, T645-18, T878-18, and T970-18. All women signed the written informed consent form. All participation in the studies was voluntary and the woman could withdraw from the study at any time.

The ethical aspects in Paper I mainly concerned whether it was safe and ethically acceptable to treat pregnant women with vaginal progesterone. Women in Paper I with a "short cervix" were invited to participate in the OPPTIMUM trial (http://www.opptimum.org.uk) an international, multicenter RCT in which women considered to be at risk were randomized to vaginal progesterone or placebo groups. Extensive human data has not shown any teratogenic effects of progestogens in the late second or third trimester[171-176]. In assisted reproductive technology the same drug has been used in the first trimester in women without any reports of increased fetal risks. Administration of progesterone in the OPPTIMUM study (22-34 GW) was therefore unlikely to have any teratogenic effect.

In the CERVIX study (Paper II) the cervical measurement was not disclosed to the women or the caregivers. The CERVIX study may therefore be considered unethical as studies from other countries have shown that a "short cervix" is a risk factor for PTD and there may be possible preventive strategies. There was no data before Papers I and II on the prevalence of a "short cervix" in Sweden, nor on the best cut-

off point in cervical length for the prediction of PTD, nor the best time to measure cervical length. A meta-analysis has shown that treatment with vaginal progesterone for women with a "short cervix" can reduce the rate of PTD[88]. However, not all studies have shown a positive effect on PTD[87,101,104,105]. It is also important to take into consideration that it is the consequences of PTD for the infant that are most important. Reducing the rate of PTD in itself is not a value and PTD may be considered as a surrogate outcome measure. Spontaneous PTD is associated with intrauterine inflammation and infection[83,177] and prevention of sPTD may cause harm by maintaining the fetus in an adverse environment. Significant reduction of short-term neonatal adverse outcome has been reported in some studies[86,88] but not in all[87,103]. There is little follow up data on infants and a long-lasting positive effect on infants has not been proved[87,178]. In Sweden, there are no national guidelines for universal screening for cervical length , or for screening of women at risk. A recommendation for screening or preventive treatment is up to the decision of the individual clinician. The participating women were informed that the measurement results would not be dísclosed to themselves or to the caregivers and not used for pregnancy management. Only if it was found during examination that membranes were bulging into the vagina, was the woman informed and referred to an obstetrician.

When evaluation of intra-observer agreement was planned in the reproducibility study, we thought a maximum of two examinations would be acceptable to women. Therefore intra-observer reproducibility was evaluated in the CLIPS study but not in the LIVE study.

4. Results and comments

4.1 Paper I

Results

In total, 2122 asymptomatic women underwent cervical length measurement using transvaginal ultrasound between 16+0 GW and 23+1 GW (mean 18+4 GW), of whom 94% were screened between 16+0 and 19+6 GW. Ninety two per cent of all cervical measurements were performed in Gothenburg and 8% in Trollhättan. Cervical length was approximately normally distributed and mean (SD) cervical length was 39.9 (6.5) mm. Eleven women (0.5%) had a cervical length ≤25 mm and 73 (3.4%) ≤30 mm. Seven out of 11 women with a cervical length ≤25 mm were included in the OPPTIMUM trial[87] and were excluded from further analysis, as well as women with indicated PTDs (n=35) and women lost to follow-up (n=19). In total, 2061 women were analysed regarding the risk of sPTD.

There were significantly more nulliparous women in the study population (923/2061) than in the non-screened population (2843/7216), 49% vs 43%, but otherwise the groups were comparable concerning maternal characteristics (Table 1, Paper I). In the study population the rate of sPTD <34 GW was 1.0% (n= 22) and <37 GW 4.2% (n= 87), in the non-screened population it was 1.0% and 3.7%, respectively. The relative risk (RR) of sPTD <34 GW was 15.8 (95% CI 1.9-134.5) if cervical length was at or below the 10th percentile (<32 mm).

The mathematically best cut-off point on the ROC curve was a cervical length of 37 mm corresponding to the 35th percentile (Figure 13) which for a prediction of sPTD <34 GW and <37 GW had a sensitivity of 59% and 53%, a PPV of 2% and 6% and an AUC of 0.689 and 0.582, respectively.

The univariable logistic regression analysis showed a significant association between cervical length and sPTB <34 GW (OR 1.78; 95% CI 1.19-2.65, p=0.005, for a decrease of cervical length by 5 mm) (Figure 14). No significant association was found for CL and sPTB <37 GW. Adjustment for relevant confounders only marginally changed the result.

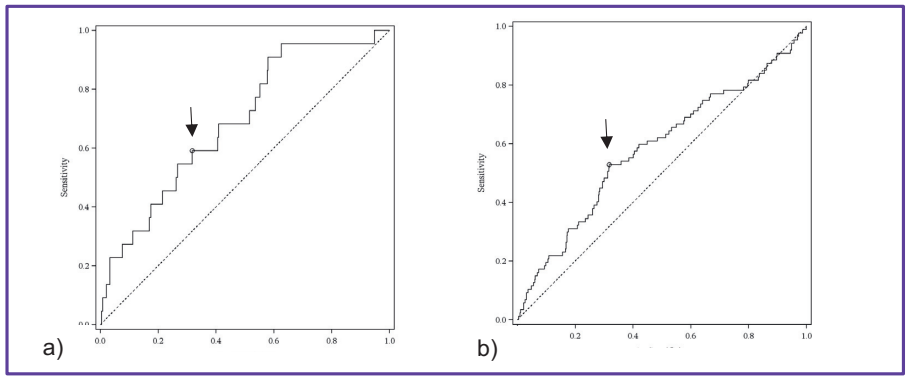


Figure 13 Receiver operating characteristics curves showing the ability of cervical length to predict spontaneous preterm delivery before 34 GW (a) and 37 GW (b). The arrows indicate the mathematically best cut-off point (cervical length 37 mm)

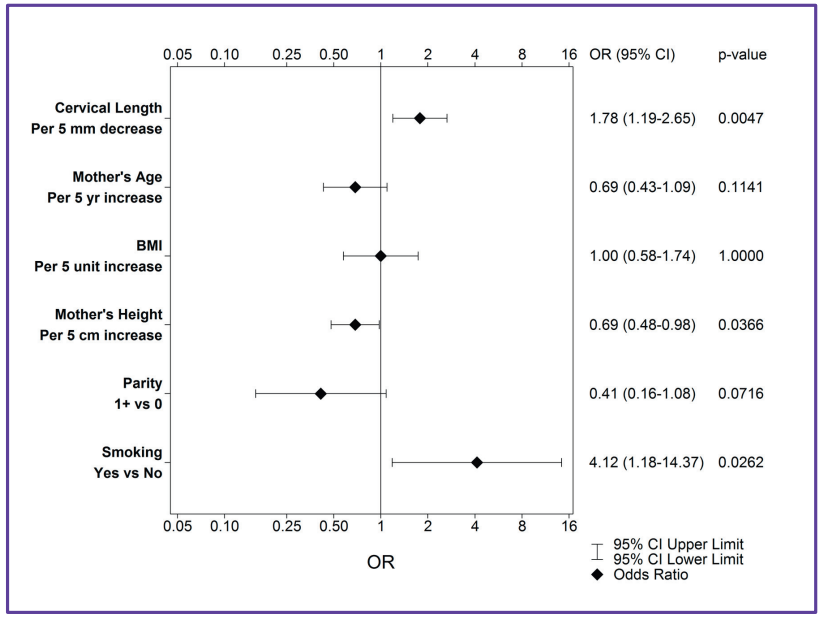


Figure 14 Results from univariable logistic regression analysis of selected maternal characteristics for the risk of preterm delivery <34 GW. For cervical length, odds ratio (OR) is the ratio for the odds of a decrease of the predictor by five units. For mother's age, BMI, mother's height, OR is the ratio for the odds of an increase of the predictor by five units. Parity (parity 0 or parity 1+) and smoking (yes or no at first antenatal visit) were categorical variables CI, confidence interval

Comments

Short cervical length and smoking were associated with sPTD <34 GW (the primary outcome), but not with sPTD <37 GW in this study. It confirmed the findings from other studies that the shorter the cervical length, the higher the risk of sPTD[64,69]. The prevalence of short cervix, if defined as ≤25 mm, was lower, 0.5% (n=11), than expected. Seven of these women with a short cervix were included in the OPPTIMUM trial and were randomised to progesterone or placebo. They delivered at term except two: one at 36+4 GW and another at 36+5 GW. Four women with a short cervix who declined participation in the OPPTIMUM trial delivered at term. Other studies have found a prevalence of short cervical length (≤25 mm) varying between 3.2% and 10.7%[64,67,90,155,179]. One study from Finland showed a prevalence of cervical length ≤25 mm of 0.3% which corresponds to our findings[69].

Measurement of cervical length had poor sensitivity and low PPV to predict sPTD. However, the results were based on only 22 women who had a sPTD ≤34 GW and we concluded that a larger study was needed before cervical length measurement could be recommended as a universal screening method.

4.2 Paper II

Results

During the study period 54 668 women in seven participating centers attended the routine pregnancy scan at 18 GW and 76% of them were assessed for eligibility. Of the eligible women, 46% declined participation leaving 11,456 recruited women for cervical length assessment. A flow chart shows the study populations and reason for exclusions (Figure 2, Paper II). The study populations consist of 11,072 women in the Cx18W population (result from cervical measurement at Cx1 and outcome data), 6288 women in the Cx21W population (result from cervical measurement at Cx2 and outcome data) and 6179 women in the Cx18W21W (result from cervical measurement at both Cx1 and Cx2 and outcome data). The No Cervix measurement population consists of 9799 women and the background population of 347,398 women. There are some differences in background and outcome data between the groups (Table 8 and Table 9). In the study populations and No Cervix population fewer women were born outside Europe at between 7% and 10%, compared with 18% in the Background population. In the study populations there were more women

with in vitro fertilization pregnancies and women with a history of conization or previous singleton sPTD (Table 8). The rate of sPTD was higher in the Background population than the study populations (Table 9).

Table 8 Background characteristics by study groups

Variable	Cx18W n=11 072	Cx21W (n=6288)	Cx18W21W n=6179	No cervix measurement (decliners) n=9799)	Swedish background population (n=347 398)
Age at delivery (years)	31.3 (28.3; 34.8)	31.3 (28.3; 34.6)	31.3 (28.3; 34.6)	30.9 (27.7; 34.3)	30.7 (27.1; 34.4)
Maternal country of birth outside European countries	1008 (9.9%)	426 (7.3%)	420 (7.3%)	854 (9.5%)	56 439 (18.2%)
BMI at first antenatal visit	23.7 (21.6; 26.8) n=9968	23.8 (21.6; 26.9) n=5730	23.8 (21.6; 26.9) n=5634	23.7 (21.5; 27.0) n=8530	23.8 (21.5; 27.1) n=322 139
Smoking	390 (3.9%)	189 (3.3%)	186 (3.3%)	375 (4.5%)	15 868 (4.7%)
IVF in current pregnancy	587 (5.5%)	359 (5.9%)	350 (5.9%)	376 (4.0%)	15 260 (4.6%)
Conization of cervix before pregnancy	654 (5.9%)	403 (6.4%)	396 (6.4%)	415 (4.2%)	14 710 (4.3%) n=341 000
Nulliparous	5223 (49.5%)	3131 (51.6%)	3084 (51.7%)	4175 (45.2%)	160 345 (47.7%)
Previous PTD (<37 GW) ≥1*	503 (4.5%)	303 (4.8%)	295 (4.8%)	308 (3.1%)	11 357 (3.3%) n=341 000
Previous singleton spontaneous PTD (<37 GW) ≥1	372 (3.4%)	223 (3.5%)	221 (3.6%)	205 (2.1%)	7825 (2.3%) n=341 000

For categorical variables n (%) is presented
For continuous variables Median (interquartile range Q1; Q3) is presented
BMI=body mass index (kg/m^2), GW=gestational weeks, IVF=in vitro fertilization, PTD=preterm delivery
* previous singleton, multiple, spontaneous and indicated PTDs are included

Table 9 Pregnancy and delivery outcome by study groups

Variable	Cx18W n=11 072	Cx21W (n=6288)	Cx18W21W n=6179	No cervix measurement (decliners) n=9799)	Swedish background population (n=347 398)
Redeemed prescription of vaginal progesterone after inclusion/18+0 GW* until delivery	18 (0.2%)	9 (0.1%)	8 (0.1%)	0 (0.0%)	823 (0.2%) n=341 000
Cerclage after inclusion/ 18+0 GW*	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	244 (0.1%) n=341 000
Preeclampsia or gestational hypertension (at delivery)	619 (5.6%)	368 (5.9%)	359 (5.8%)	534 (5.4%)	16 890 (4.9%)
Induction of labour	1946 (17.6%)	1090 (17.3%)	1066 (17.3%)	1591 (16.2%)	59 259 (17.1%)
Caesarean delivery	1704 (15.5%)	944 (15.1%)	923 (15.0%)	1475 (15.2%)	60 088 (17.3%)
Late miscarriage† (before 22+0 GW)	8 (0.1%)	0 (0.0%)	0 (0.0%)	7 (0.1%)	NA
Spontaneous late miscarriage (before 22+0 GW)	7 (0.1%)	0 (0.0%)	0 (0.0%)	4 (0.0%)	NA
PTD 22+0-32+6 GW	109 (1.0%)	53 (0.8%)	52 (0.8%)	72 (0.7%)	3765 (1.1%)
PTD 22+0-36+6<37 GW	577 (5.2%)	321 (5.1%)	313 (5.1%)	412 (4.2%)	18 101 (5.2%)
Spontaneous PTD 22+0-27+6 GW	15 (0.1%)	3 (0.0%)	3 (0.0%)	18 (0.2%)	993 (0.3%)
Spontaneous PTD 22+0-32+6 GW	56 (0.5%)	26 (0.4%)	26 (0.4%)	42 (0.4%)	3302 (1.0%)
Spontaneous PTD 22+0-36+6 GW	410 (3.7%)	225 (3.6%)	220 (3.6%)	293 (3.0%)	15 357 (4.4%)

For categorical variables n (%) is presented.
GW=gestational weeks, PTD=preterm delivery
*18+0 GW for Swedish background population
†includes induction of labour for missed abortion

The ability to predict sPTD between <28 and <37 GW was comparable for the nine measurements of cervical length (minimum [shortest], mean or maximum distance of A-B, A-C, A-B + B-C) (Figure 15), whether the isthmus was present or not. Therefore, only the results showing the minimum (shortest) endocervical length are presented below. This is the measurement most widely used by others [63,88,154]. Results for min A-C and mean A-C are presented in Appendix in Paper II.

The median gestational age at Cx1 was 19+0 GW (interquartile range [IQR] 18+3 to 19+3) and at Cx2 23+0 GW (IQR 22+4 to 23+3), the median interval between Cx1 and Cx2 was 28 days (IQR 24-31). Median cervical length was 36.0 mm (min 3.0, max 60) at Cx1 and 36.0 mm (min 4, max 60) at Cx2. Isthmus was present in 23.3% of the women at Cx1 and 8.9% at Cx2. The prevalence of different cut-offs of endocervical length is presented in Table 10.

The probability of continued pregnancy without delivery at different gestational ages before 37+0 GW in relation to cervical length at Cx1 and Cx2 is shown in Kaplan-Meier plots in Figure 16a. Figure 16b with indicated PTD is censored.

Table 10 Prevalence of different cut-offs of shortest endocervical length at Cx1 and Cx2

Cervical length cut-off	Cx1 (18+0-20+6 GW) n=11 072		Cx2 (21+0-23+6 GW) n=6288	
	n	%	n	%
≤10 mm	7	0.06	6	0.10
≤15 mm	15	0.14	14	0.22
≤20 mm	67	0.61	71	1.13
≤25 mm	**441**	**3.98**	**274**	**4.36**
≤30 mm	2175	19.6	1166	18.6

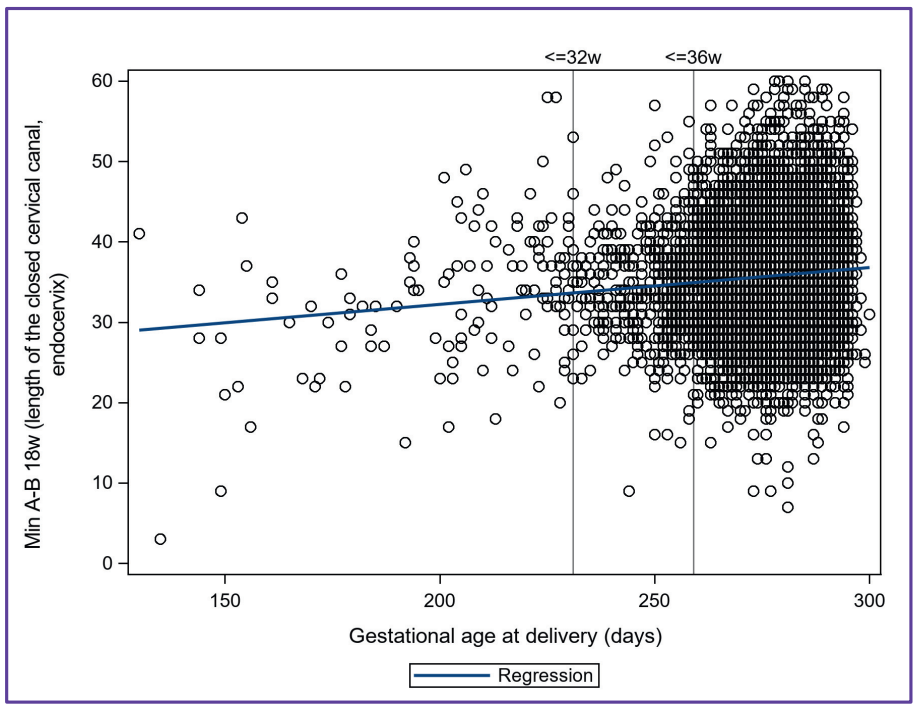


Figure 15 shows a scatter plot of all measurements of shortest endocervical length at Cx1 vs time of delivery (days)

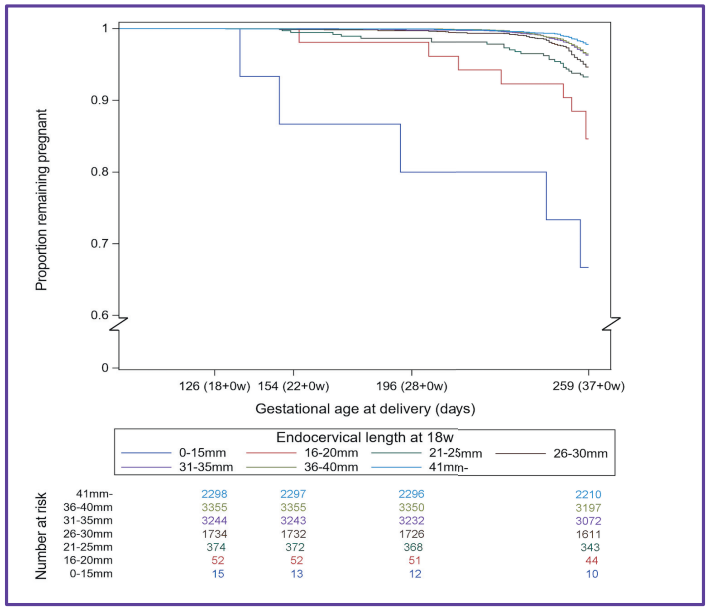


Figure 16a

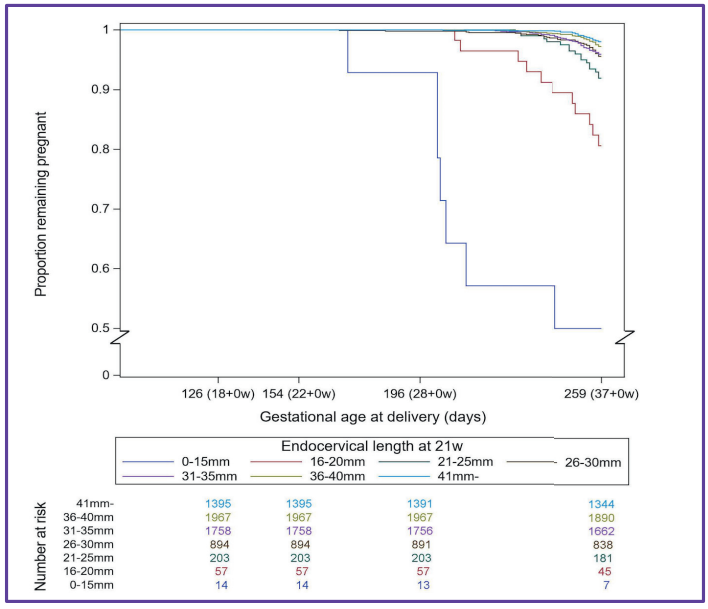


Figure 16b

Figure 16 Kaplan-Meier plots to illustrate the probability of continued pregnancy at different gestational ages before 37+0 GW in relation to cervical length at 18+0-20+6 gestational weeks (Cx1) in Figure 16a, and at 21+0-23+6 GW (Cx2) in Figure 16b. Women with indicated PTD being censored. w= weeks

The ability of endocervical length to predict PTD between 22+0 and 32+6 GW (primary outcome) at Cx1 and Cx2 was poor (AUC 0.57 and AUC 0.59, respectively). The ability of endocervical length to predict sPTD <33+0 GW, including late miscarriage (secondary outcome), at Cx1 was better (AUC 0.74) if the isthmus was absent than if the isthmus was present (AUC 0.57). To study which of the measurements at Cx1 or Cx2 was best to predict sPTD, the Cervix18W21W population was analysed. Endocervical length at Cx2 was better than Cx1 at predicting sPTD <33 GW (AUC 0.76 and 0.65, respectively) and also better than changes in endocervical length between Cx1 and Cx2 (AUC 0.67).

In Table 11 the ability of endocervical length to predict sPTD <28, <33 and <37 GW at Cx1 and Cx2 is shown. In the Cx18W population, the 25 mm endocervical length cut-off identified 27% (17/63) of the women with sPTD <33 gestational weeks with NNH 25, and NNS 651. Using the mathematically best cut-off (29 mm) the corresponding numbers were 43% (27/63), 60 and 410.

There were seven late spontaneous miscarriages in the Cx18W population, three of which occurred after PPROM. They occurred between 3 and 16 days after Cx1 measurement. Cervical length varied from 3 mm to 41 mm (four were ≤25 mm). In Cx21W population there were no women with late miscarriages.

Table 11 Discriminative ability of endocervical length measured at 18+0 to 20+6 GW (Cx1); Cx18W population (n=11,072), and at 21+0 to 23+6 GW (Cx2); Cx21W population (n=6288), with regard to predicting spontaneous preterm delivery

Cervical length at Cx1									
			0-25 mm (n=441)			**Best cut-off***			
sPTD	**No. sPTD (%)**	**AUC**	**Sens (%)**	**PPV (%)**	**NNH/ NNS**	**Sens (%)**	**PPV (%)**	**NNH/ NNS**	**mm**
<28 GW	22 (0.20%)	0.83	9/22 (40.9%)	9/ 441 (2.0%)	48/ 1230	18/22 (81.8%)	18/3383 (0.5%)	187/615	32 (n=3383)
<33 GW	63 (0.57%)	0.68	17/63 (27.0%)	17/441 (3.9%)	25/ 651	27/63 (42.9%)	27/1638 (1.7%)	60/410	29 (n=1638)
<33 GW†	56 (0.51%)	0.66	13/56 (23.2%)	13/437 (3.0%)	33/ 851	38/56 (67.9%)	38/4682 (0.8%)	122/291	34 (n=4682)
<37 GW	417 (3.77%)	0.60	38/417 (9.1%)	38/441 (8.6%)	11/ 291	207/417 (49.6%)	207/4013 (5.2 %)	18/54	33 (n=4013)
Cervical length at Cx2									
			0-25 mm (n=274)			**Best cut-off***			
sPTD	**No. sPTD (%)**	**AUC**	**Sens (%)**	**PPV (%)**	**NNH/ NNS**	**Sens (%)**	**PPV (%)**	**NNH/ NNS**	**mm**
<28 GW	3 (0.05%)	0.96	1/3 (33.3%)	1/274 (0.36%)	273/6288	3/3 (100%)	3/510 (0.59%)	169/2096	27 (n=510)
<33 GW	26 (0.41%)	0.76	10/26 (38.5%)	10/274 (3.65%)	27/629	**14/26 (53.9%)**	**14/510 (2.8%)**	**35/449**	**27 (n=510)**
<37 GW	225 (3.58%)	0.63	34/225 (15.1%)	34/274 (12.41%)	7/185	144/225 (64.0%)	144/2926 (4.9%)	19/44	35 (n=2926)

AUC=area under receiver operating characteristic curve, CI=confidence interval, GW=gestational weeks, NNH=number needed to harm, i.e. number of false positive results per one true positive test result, NNS=number needed to screen, i.e. number of women needed to screen to detect one true positive test result, No=number of, PPV=positive predictive value, sens=sensitivity, sPTD=spontaneous preterm delivery,
*the mathematically best cut-off being the one corresponding to the point on the receiver operating characteristic (ROC) curve situated farthest from the reference line.
†n=11 064 as denominator (late miscarriage excluded).

Comments

There was an association between cervical length measured at Cx1 and Cx2, and PTD. The shorter the cervical length, the higher the risk for PTD, as shown in the Kaplan Meier plots. Cervical length in the second trimester was poor at predicting PTD, but better at predicting sPTD, especially for spontaneous, extreme and very early PTD. Endocervical length at Cx2 discriminated better than endocervical length at Cx1 between women who delivered spontaneously preterm and those who did not. The mathematically best cut-off value for short cervical length measured at Cx1 or Cx2 varied between 27 mm and 35 mm to predict sPTD <28, <33, <37 GW and with an AUC, between 0.60 and 0.96. A cervical length cut-off of 27 mm at Cx2 measurement (8% [510/6288]) identified 54% (14/26) of sPTD <33 GW, with 35

false positives per one true positive test result. Four hundred and forty-nine women needed to be screened to identify one woman with a sPTD <33 GW.

The prevalence of cervical length ≤25 mm was much higher than in our previous study, 4.0 % vs 0.5% (Paper I). This did not correspond to a higher rate of sPTD in Paper II. The rate of sPTD <37 GW was 4.2% in Paper I (87/2061) and 3.7% (410/11072) in Paper II. The corresponding figures for sPTD <34 GW were 1.1% (22/2061) and 0.8% (87/11 072), respectively. This study was a multicenter study and cervical measurements were assessed by 25 midwives at seven centers, while the first was mostly done in one center with measurements performed by eight midwives. In the first study (Paper I) the median cervical length was four mm longer than in this study (39.9 mm vs 36.0 mm at both Cx1 and Cx2). This is discussed further in Chapter 5.5 p.81-84.

The total rate of PTD was comparable in the study populations and the Background population. However, the rate of sPTD was higher in the Background population (sPTD <37 GW: 4.4% vs 3.6-3.7%). The differences in rates of sPTD may be explained by different data sources and definitions. For categorization of PTD into spontaneous or indicated PTD the medical records of women with PTD in the study populations were manually scrutinized, but for the Background population only register data was used.

4.3 Paper III

Results

A summary of background data of the study populations in the LIVE and the CLIPS studies is shown in Table 12.

The LIVE study

Inter-observer continuous variables

Each pair of examiners (n = 7) assessed between 24 and 30 women each.The median length of the shortest endocervical length (n =198) was 35.0 mm (range 20.0 to 58.0 mm). The intraclass correlation coefficient (ICC) (reliability) varied between 0.31 and 0.91 for the seven examiner pairs. For one examiner pair, ICC was high and limits of agreement (LoA) was narrow; for two pairs of examiners ICC was low and

LoA wide (Table 13 and Figure 17). Agreement and reliability were better if both examiners in a pair agreed that isthmus was not present. (Figure 17).

Table 12 Background data for the LIVE and the CLIPS studies

Maternal characteristics	**Live** (n=198)	**Clips** (n=93)
Maternal age	32 (21; 45)	31 (22; 41)
BMI	24 (17; 41)	24 (18; 42)
Nulliparous	84 (42%)	47 (51%)
Ethnicity white other	 168 (85%) 30 (16%)	 82 (88%) 11 (12%)
History of late miscarriage or preterm delivery	16 (8%)	5 (5%)
History of cervical conization	15 (7.6%)	10 (10.8%)
Gestational age when CL was measured, weeks, days	19+4 (18+0; 23+6)	19+2 (18+0; 23+4)
For categorical variables n (%) is presented. For continuous variables Median (Min; Max) BMI= body mass index, CL= cervical length		

Table 13 The LIVE study: Agreement and reliability of seven pairs of examiners with regard to shortest endocervical length

		Difference in mean endocervical length (95% CI of mean difference), mm	**Limits of agreement, mm**	**ICC**
Best	Pair F	0.33 (-0.61; 1.28)	-4.06 to 4.72	0.91
	Pair A	0.97 (-0.26; 2.19)	-5.35 to 7.28	0.76
	Pair B	1.37 (-0.02; 2.7)	-5.92 to 8.65	0.63
	Pair E	0.14 (-1.31; 1.59)	-7.34 to 7.61	0.70
	Pair G	0.86 (-0.57; 2.29)	-6.37 to 8.09	0.79
Poorest	Pair C	0.73 (-1.64; 3.10)	-11.70 to 13.17	0.58
	Pair D	0.14 (-2.08; 2.37)	-11.11 to 11.39	0.31
Limits of Agreements (LoA): 1.96 standard deviations on either side of the mean We expect 95% of differences between future measurements by two examiners in a pair to fall within the LoA [180]. **Intraclass correlation coefficient** (ICC): ANOVA, two-way random model, absolute agreement as an estimate of reliability in general.				

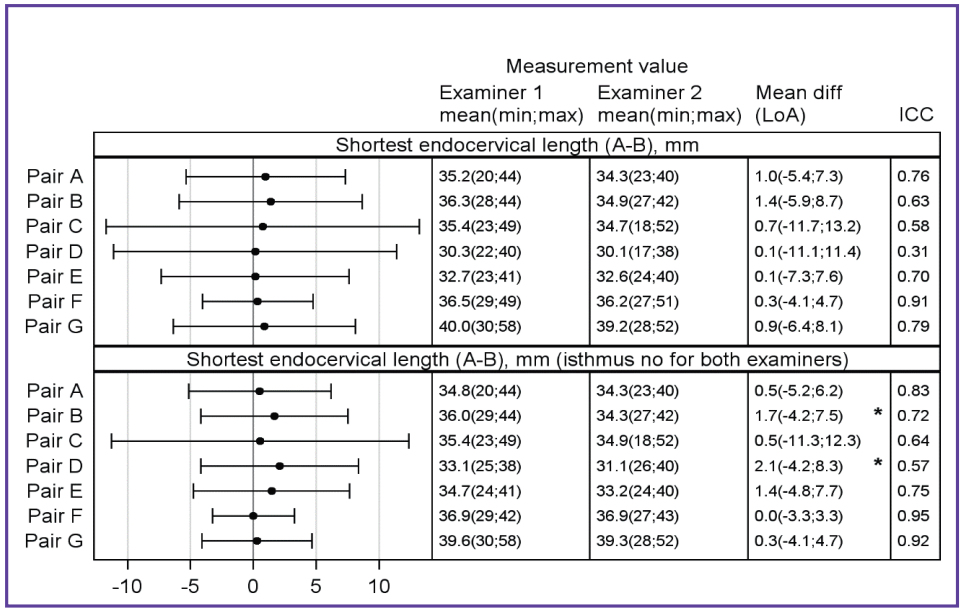


Figure 17 LIVE study. Forest plots showing mean difference (dot) and Limits of Agreements, LoA, (lines) for seven pairs of examiners taking measurements of shortest endocervical length, and for the shortest endocervical length when both examiners agreed on absence of isthmus. All mean differences in mm are shown as positive differences with LoA adjusted accordingly. Asterisks denote systematic differences between the two examiners. The measurement results and the intraclass correlation coefficient (ICC) for each pair is also shown.
Examiner 1 is the examiner with the highest mean value of the studied variable. Examiner 2 is the examiner with the lowest mean value of the studied variable.

The CLIPS study

Intra-observer continuous variables

Sixteen raters assessed 93 video clips twice with at least two months between the assessments. Seven video clips of the original 100 were excluded because some women contributed two clips.The first video clips were included in the CLIPS study. Median length of shortest endocervix (n=1485) at first assessment was 35.0 (min 10 mm; max 53 mm). Median mean difference between two measurements (intra-rater) (n = 16 raters) in endocervical length was -0.15 mm (range -1.48 to 1.27 mm) (second measurement subtracted from first), median IISD 2.14 mm (range 1.40 to 3.46 mm), median repeatibility 5.93 mm (range 3.88 to 9.58). ICC for endocervical length was 0.84 (range 0.66 to 0.94). The corresponding values are shown in Table 14 for each of the 16 raters.

Figure 18 illustrates with a Bland-Altman plot the differences between each rater´s measurements and the mean of all 16 raters` measurements of endocervical length.

Table 14 CLIPS study. Intra-observer differences, measurement error (intra-individual standard deviation [IISD]), repeatability and reliability with regard to measurements of shortest endocervical length for each of 16 raters (n=93 video clips for each rater, except rater 7 [n=92], rater 12 [n=91])

	Endocervical length			
Unique rater number	**Difference,** mm Mean (SD), Min; Max	**Measurement error IISD**, mm	**Repeatability,** mm	**Reliability ICC**
Rater 1	-0.01 (2.77), -10.0; 11.0	1.95	5.40	0.85
Rater 2	-0.04 (2.60), -13.0; 8.0	1.83	5.07	0.87
Rater 3	1.27 (2.83), -8.0; 9.0	2.18	6.04	0.80
Rater 4	-0.74 (2.77), -15.0; 6.0	2.01	5.57	0.87
Rater 5	-0.63 (2.31), -12.0; 6.0	1.68	4.65	0.88
Rater 6	-0.03 (3.32), -14.0; 16.0	2.34	6.48	0.86
Rater 7	0.61 (4.88), -12.0; 11.0	3.46	9.58	0.73*
Rater 8	0.43 (3.88), -19.0; 15.0	2.75	7.62	0.80
Rater 9	0.52 (3.14), -10.0; 11.0	2.24	6.20	0.80
Rater 10	0.45 (3.97), -6.0; 24.0	2.81	7.78	0.69*
Rater 11	-0.25 (3.44), -15.0; 12.0	2.43	6.73	0.79
Rater 12	-1.32 (2.68), -11.0; 8.0	2.10	5.82	0.84
Rater 13	-0.51 (2.93), -11.0; 10.0	2.09	5.79	0.83
Rater 14	-1.33 (1.58), -7.0; 3.0	1.46	4.04	0.92†
Rater 15	-1.48 (4.68), -17.0; 12.0	3.45	9.56	0.66*
Rater 16	-0.30 (1.96), -7.0; 8.0	1.40	3.88	0.94†
*denotes the three poorest raters † denotes the two best raters				

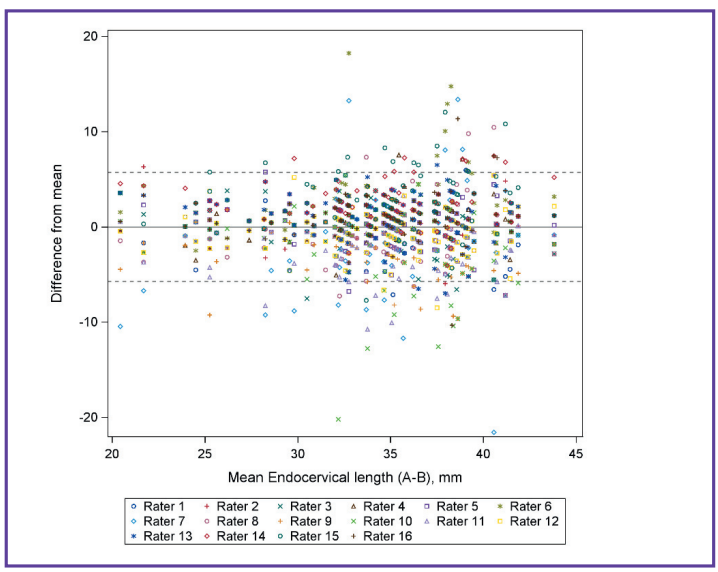


Figure 18 CLIPS-study. Plot of differences between each rater´s measurement (Y-axis in mm) against the mean of all 16 raters´ measurements (X-axis) of shortest endocervical length. Each rater is represented by one symbol, but symbols may be superimposed. The dotted horizontal lines represent the 95% Limits of Agreement (LoA) with the mean[181]. The lower LoA is -5.7 mm and the upper LoA is 5.7 mm. Rater 7, 10 and 15 seem to deviate from the others

LIVE and the CLIPS study

Agreement on and reliability of categorical variables

Agreement on and reliability of the shortest cervical length ≤ 25 mm (present or not) was studied in the seven pairs in the LIVE study. In the CLIPS study, agreement on and reliability of the shortest cervical length ≤25 mm (present or not) was studied in the first analysis in 120 rater pairs (constructed from 16 raters) and between first and second analysis of 16 raters in the CLIPS study (Table 15). In the LIVE study four out of seven pairs recorded the shortest cervical length ≤25 mm, measured by at least one of the two examiners.

Negative agreement between examiners and within raters was good, while positive agreement was quite poor.

Similarly, agreement and reliability about whether isthmus was present or not in the LIVE study and the CLIPS study is shown in Table 16.

As for cervical length ≤25 mm negative agreement between examiners and within raters was generally good, while positive agreement was quite poor.

Table 15 Inter-observer (LIVE, CLIPS) and intra-observer (CLIPS) agreement and reliability for cervical length ≤25

	Agreement			**Reliability**
	Total agreement (%)	**Positive agreement** (%)	**Negative agreement** (%)	**Cohen's kappa**
LIVE (inter-observer), 7 examiner pairs	100.0 (78.6; 100)	32.5 (0; 100)	100 (87.5; 100)	0.26 (-0.05; 1.00)
CLIPS (inter-observer), 120 rater pairs (16 raters). Analysis 1	94.6 (84.9; 98.9)	58.8 (13.3; 92.3)	97.1 (92.3; 99.5)	0.56 (0.12; 0.92)
CLIPS (intra-observer), 16 raters, Analysis 1 and Analysis 2, >2 months apart	95.2 (87.1; 98.9)	69.7 (28.6; 93.3)	97.4 (93.3; 99.4)	0.68 (0.27; 0.93)
Distribution over all seven examiner pairs, 120 rater pairs and 16 raters, respectively: Median (Min; Max) *Positive agreement = (2x Yes both Examiners)/(Yes Examiner 1 + Yes Examiner 2)[182] †Negative agreement = (2x No both Examiners) /(No Examiner 1 + No Examiner 2)[182]				

Table 16 Inter-observer (LIVE, CLIPS) and intra-observer (CLIPS) agreement and reliability on presence or absence of isthmus

	Agreement			**Reliability**
	Total agreement(%)	**Positive agreement** (%)	**Negative agreement** (%)	**Cohen's kappa**
LIVE (inter-observer), 7 examiner pairs	93.3 (82.8; 96.4)	70.6 (0.0; 94.7)	95.8 (87.8; 98.2)	0.69 (0.27-0.91)
CLIPS (inter-observer), 120 rater pairs (16 raters). Analysis 1	81.7 (59.1; 93.5)	52.0 (22.2; 92.7)	88.9 (71.3; 96.5)	0.42 (0.12; 0.87)
CLIPS (intra-observer), 16 raters, Analysis 1 and Analysis 2, >2 months apart	89.2 (74.2; 94.6)	70.7 (40.0; 90.2)	93.4 (80.3; 97.1)	0.66 (0.36; 0.87)
Distribution over all seven examiner pairs, 120 rater pairs and 16 raters, respectively: Median (Min; Max) *Positive agreement = (2x positive isthmus both Examiners)/(positive isthmus Examiner 1 + positive isthmus Examiner 2)[182] †Negative agreement = (2x negative isthmus both Examiners) /(negative isthmus Examiner 1 + negative isthmus Examiner 2)[182]				

Comments

We found substantial differences between the seven pairs of examiners in the LIVE study regarding inter-observer agreement and reliability, and substantial differences between the 16 raters in the CLIPS study regarding intra-observer measurement error, repeatability and reliability. The differences in results between examiner pairs and between raters are likely to be explained by differences in the skill and carefulness of the examiners, and also perhaps by differences in stringency of local supervision.

In the LIVE study inter-observer differences (LoA) in endocervical length was < ±5 mm in one pair, between ± 5 and 9 mm in four pairs and approximately ±11 mm in two pairs. When both examiners agreed that isthmus was absent, then two out of seven pairs had LoA < ±5 mm when the endocervical length was measured (Figure 17). In the LIVE study the width of LoA varied between 9 mm and 25 mm for the seven examiner pairs and in the CLIPS study between 10 mm and 25 mm for the 120 pairs of raters. These are similar to the study by França et al who reported a width of LoA between 12 mm and 23 mm)[165].
In the CLIPS study, repeatability varied between 3.9 mm and 9.6 mm among the 16 raters.

Inter-observer agreement and reliability differed markedly between different observer pairs in both LIVE and CLIPS studies for the two categorical variables, isthmus present or not and cervical length ≤25 mm or not. It was easier for pairs to

agree that the isthmus was absent than it was for them to agree it was present. If an isthmus is present the recognition of the inner cervical os can be difficult and endocervical length may be measured as longer if an isthmus is incorrectly included in the measurement. Only three other studies have described how to measure cervical length when the isthmus is present, and awareness of its existence may not be widespread[142,183,184].

A cut-off of cervical length ≤25 mm is often used to identify women at risk of PTD. In this study negative agreement on cervical length ≤25 mm was generally good, while positive agreement varied considerably. The results from the LIVE study are difficult to interpret because there were very few women with cervical length ≤25 mm. Our results show that intra-observer and inter-observer agreement and reliabilty with regard to cervical length ≤25 mm are not very reliable. If used for clinical decisions regarding preventive treatment and hospitalization, or recruitment to RCTs, re-assessment by a second examiner may be needeed.

5. General discussion

5.1 Study design

One of the goals of transvaginal cervical length measurement is to find women at risk of PTD. We have performed two observational studies (Paper I and II) and two reproducibility studies (Paper III). The first observational study (Paper I) may be considered as a PILOT study where women were also recruited to an RCT and therefore cervical measurements were not blind. The primary outcome differed between the studies; spontaneous PTD <34 GW versus any PTD <33 GW. In the second observational study (Paper II) the results of the cervical measurement were blind both for the woman and the care providers.

Sample size calculation in the second observational study was based on the aim of getting at least 100 women with PTD <33 GW which could give a reasonable confidence interval when it came to sensitivity to predicting PTD. The study population then needed to include 11,000 women. (www.socialstyrelsen.se/publikationer2018/2018-1-6). We reached the goal with almost 11,500 women included and 109 women with PTD between 22+0 and 32+6 GW, rising to 117 women when late miscarriages were included. A more optimal primary outcome would have been sPTD <33 GW, however the study population would then have needed to include almost twice the number of participants. The distinction between spontaneous and indicated PTD is not always easy and therefore we found it relevant to include all PTDs.

Our first observational study indicated a need for a larger observational study as only 22 women had sPTD <34 GW (1.1%). Risk calculations are uncertain and there are often wide confidence intervals when studies are based on few cases, which was the case in the early gestational weeks. Our study population in Paper II is more than three times as large as the largest blind study (n=3694 women)[69] and comparable with the largest non-blind study (n=11,943 women)[65] that we found in a systematic literature review of observational studies of cervical measurements on asymptomatic women (Table 4 and 5, p 40-43).

In our systematic literature, 13 studies showing reproducibility data on cervical measurement were identified (presented in Appendix pages 3 to 7 in Paper

III)[54,60,165,166,184-192]. All 13 studies were single center studies. Only two relevant studies were included[60,165]. Reproducibility of measurements was studied by only a few examiners. Our reproducibility study was performed by several specially trained midwives (7 pairs of examiners in the LIVE study and 16 raters in the CLIPS study) at seven centers, making the results more generalizable.

5.2 Methodological aspects - Internal validity (information bias, selection bias and confounders)

Internal validity

Information bias

In Paper II quality controls of values of cervical measurement and dates in eCRF were performed before closing the data set. Measurement values in eCRF from a sample of twenty random cases (9-19 values per case) per center were checked against documented images. A maximum error rate was set at ≤1%. Five out of seven centers were approved at the first attempt, the rest at the second attempt.

All extreme values and all endocervical lengths ≤7 mm were identified in eCRF and checked against the documented images.

Incorrect classification of PTD into spontaneous or indicated PTD may have occurred in the Pregnancy Register and the Medical Birth Register. For the Background population in Paper II, aggregated registry data was used for classification of PTD. e.g. ICD 10 codes and procedures registered at discharge from the hospital after delivery. For the study populations, both individual registry data and medical records were used for classification of PTD. Scrutinizing the medical records changed the classification of PTD from registry data in some cases i.e. if a PPROM code was missing and induction of labour was performed, the delivery was incorrectly classified as an indicated PTD by using registry data. The validity of other registry data is commented on in Chapter 3.3 p.55.

In Paper III the data was collected in a study protocol in the LIVE study (Paper III) from both examiners in the seven pairs. All data transferred from the paper protocol to Excel was manually checked by two controllers (PK and UBW).

Selection bias

The study population should be representative of the population where the results are to be applied. To investigate whether there was any selection bias, we compared the background characteristics between the study populations and control groups. The control group in Paper I consisted of those who were not screened. In Paper II, one control group consisted of decliners and another of the Swedish Background population. None of the six blinded observational studies in the systematic literature review (see Chapter 1.5 p. 39) has used a background population for comparison[64,68,69,76,155,156].The comparison of the study populations and the control groups showed some differences regarding background characteristics with more high risk women in the study population. However, there was no evidence of substantial selection bias.

Recruitments in Paper II were mostly from university hospitals and not from private caregivers. Gothenburg recruited 35% of the women, Malmö and Lund 28%, Falun 16%, Stockholm 13% and Örebro 9%. Gothenburg and Malmö recruited during the whole study period while the other centers started recruitment later.

Confounding

Confounding factors are variables that correlate with both the outcome and the explanatory variable and may confuse the effect if not accounted for in the analysis. In observational studies there is always a risk of potential residual confounding by unknown confounders. In Paper I for example, smoking was found to be a potential confounder. In Paper II we found in an univariable analysis that ethnicity, height, level of education, infertility, parity and previous PTD were factors associated with PTD (data not shown in Paper II).

External validity

External validity refers to whether the conclusions drawn from a study are applicable to other populations. Is is reasonable to apply the result from our studies to the entire population in Sweden or to other populations in other countries? As previously mentioned, the prevalence of PTD and a short cervix differs between countries and between multicenter studies. It may affect how easily the results can be generalized to apply to another population. The results from the CERVIX study (Paper II) have good internal validity and substantial selection bias was excluded by comparing the

study population with the non-participants and the Background population. Our study cohort was fairly representative of the total underlying population but had more women born outside Europe. However, we believe that our results are generalizable to the general Swedish population.

The reproducibility study is the first multicenter study to assess reproducibility of transvaginal ultrasound measurements of cervical length, a study which involved a large number of well trained and certified examiners. We believe that our results are generalizable and reflect clinical reality.

5.3 Strengths

The main strength in Paper I was that the study population was compared with a control group that were not screened. All examiners were certified and all the cervical lengths ≤25 mm were also re-assessed.

The CERVIX study (Paper II) has several strengths: the study size, blinding of cervical measurements for participants and staff, a minimal loss to follow-up, careful description of the measurement technique, certification of the midwives, and rigorous quality controls of cervical measurements. Furthermore, the study included two control groups for the comparison of background data and outcome data, for the evaluation of selection bias and generealizability of results. The goal for the primary outcome, at least 100 PTDs <33 GW, was achieved. Study data (eCRF data) had good validity as well as all register data. A Statistical Analysis Plan was made in advance before the data set was locked. The study flow was described in detail. All three cervical measurements were registered and analysed without predefining which type of measurement would give the best prediction; minimum, maximum or mean. Outcome data included information on late miscarriage which was manually collected, because data was not available in registers.

The reproducibility study is a multicenter study with a large number of examiners/raters, who carried out the measurements in the LIVE and CLIPS studies. This is the first study to evaluate agreement on presence of isthmus and cervical length ≤25 mm. Quality control and certification was the same as in the CERVIX study.

5.4 Limitations

The main limitation in Paper I was that the sample size for cervical length measurements was low for the prediction of sPTD. Another limitation in Paper I was the low participation rate (22% of all eligible women) and that 92% of the study was performed in one center. Furthermore, the isthmus was not measured separately and it is possible that the isthmus was measured as a part of the endocervix, which may have contributed to the almost 4 mm longer median cervical length than in the CERVIX study. Finally, the study in Paper I was not blind for the women and for the health care providers, and thus interventions to prevent PTD were possible .

One limitation in the CERVIX study (Paper II) is that not all women were assessed for eligibility (76%). Almost half of eligible women (46%) declined the transvaginal cervical measurement but consented to data collection. Only 56% of women participating in the Cx1 measurement came back for Cx2 measurement. Despite translation of the study information into eight languages, the study population had fewer women born outside Europe, and language problems were the most common reason for ineligibility. Because the sample size was calculated for PTD <33 GW and not for sPTD <33 GW, our estimates regarding the sensitivity of endocervical length for the detection of sPTD <33 GW are less precise. Our secondary outcomes sPTD are more clinically relevant because sPTD is potentially preventable[88].

In the reproducibility study (Paper III), agreement on cervical length ≤25 mm was evaluated both in LIVE and CLIPS studies. Inter-observer agreement from the LIVE study is based on data from only four pairs, and findings from so few cases per pair of examiners must be interpreted critically. The inter-observer agreement of cervical length ≤25 mm was analyzed by the 120 pairs constructed from the 16 examiners in the CLIPS study. A limitation in the LIVE study is that intra-observer repeatability was not evaluated. It was doubtful whether the women would accept two cervical assessments by two examiners (i.e four examinations) in the same session. Intra-observer repeatibility and agreement are studied in the CLIPS study instead, but they are not completely generalizable to a live situation.

5.5 Difference in prevalence of short cervix

The prevalence of cervical length ≤25 mm was 0.5% in our first study in Paper I and 4.0% in the CERVIX study (Paper II). How could this difference be explained? The

sample size was over fivefold in the CERVIX study compared to the study in Paper I, but overall the study populations were comparable concerning maternal characteristics. However, there were slightly more high-risk women in the first study than in the CERVIX study, 9% compared with 4.5% to 4.8% of women with a previous PTD. The measurements of cervical length differed in some points. In the first study the isthmus was not measured separately and the certification and training of medical personnel did not emphasize the presence or absence of the isthmus. Furthermore, in the first study the cut-off of cervical length at 25 mm was applied in order to find women suitable for recruitment to an RCT (OPPTIMUM trial[87], and all cervical lengths ≤25 mm were re-assessed by a doctor. In the first study 92% of the measurements were performed in one center, whereas the CERVIX study was a multicenter study performed in seven centers. The CERVIX study had no set cut-off value for a short cervix, the measurement values were blind for the women and care-givers, and cervical length was not used for clinical management.

The median cervical length differed 4 mm between the studies, 39 mm as opposed to 36 mm. With a lower median cervical length, the distribution of cervical lengths moves to the left in the Gaussian (normal) distribution curve. The distribution of cervical lengths is shown in Figure 19 where the shortest endocervical length (A to B) is shown for the CERVIX study and the PILOT study. The distribution of the CERVIX study has a more normal distribution than the PILOT study, but the form of the curves is similar.

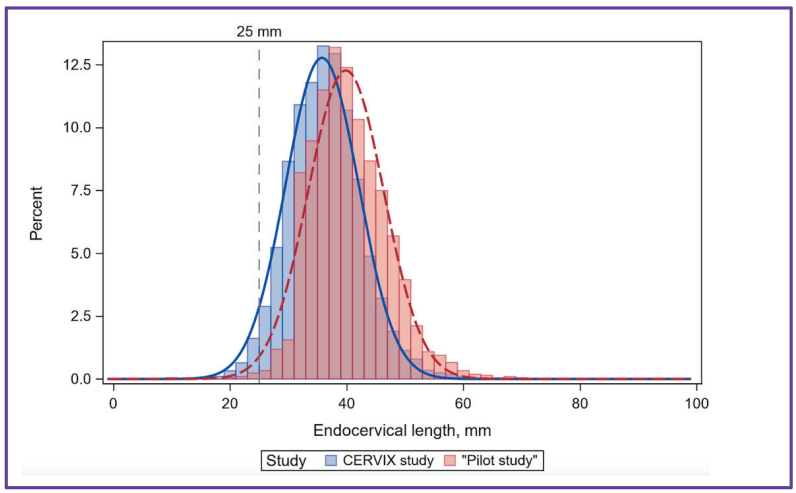


Figure 19 Distribution of the endocervical length in the CERVIX and PILOT study. Cut-off at 25 mm is shown with a dotted line

Is it possible that the isthmus is included as a part of the endocervical length and therefore the median length is longer in the PILOT study? When the distribution of endocervical length in the PILOT study is compared to the distance A to C in the CERVIX study the distributions are more similar, Figure 20.

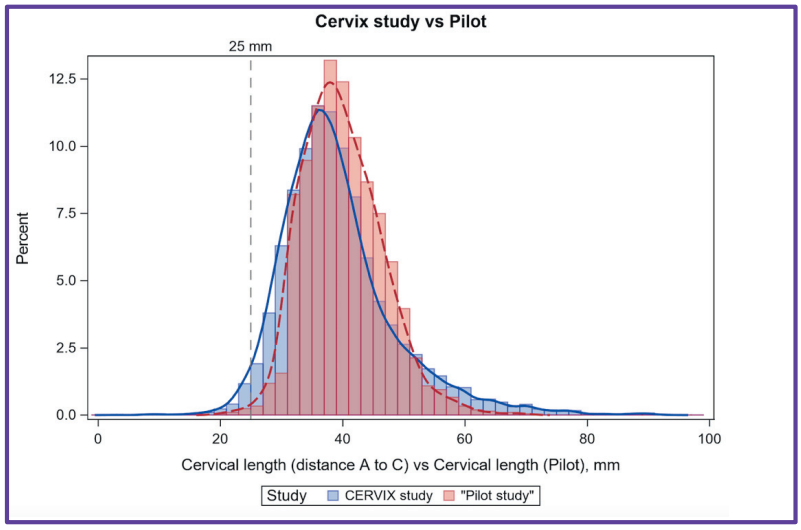


Figure 20 Distribution of the distance A to C in the CERVIX study and endocervical length in the PILOT study. Cut-off at 25 mm is shown with dotted line.

Of all cervical measurements, 35% were performed in Gothenburg, 19% in Malmö, 16% in Falun, 9% in Örebro, 9% in Lund, 8% in Solna (Stockholm) and 5% in Huddinge (Stockholm). The median endocervical length in these centers varied from 34.0 mm to 38.0 mm. In Gothenburg the median endocervical length was 35.0 mm. Variations between pairs of examiners could also be seen in the REPRODUCIBILITY study. Local supervision, personal skills and number of examiners per center can contribute to these differences.

The distribution of endocervical length is shown in the Figure 21

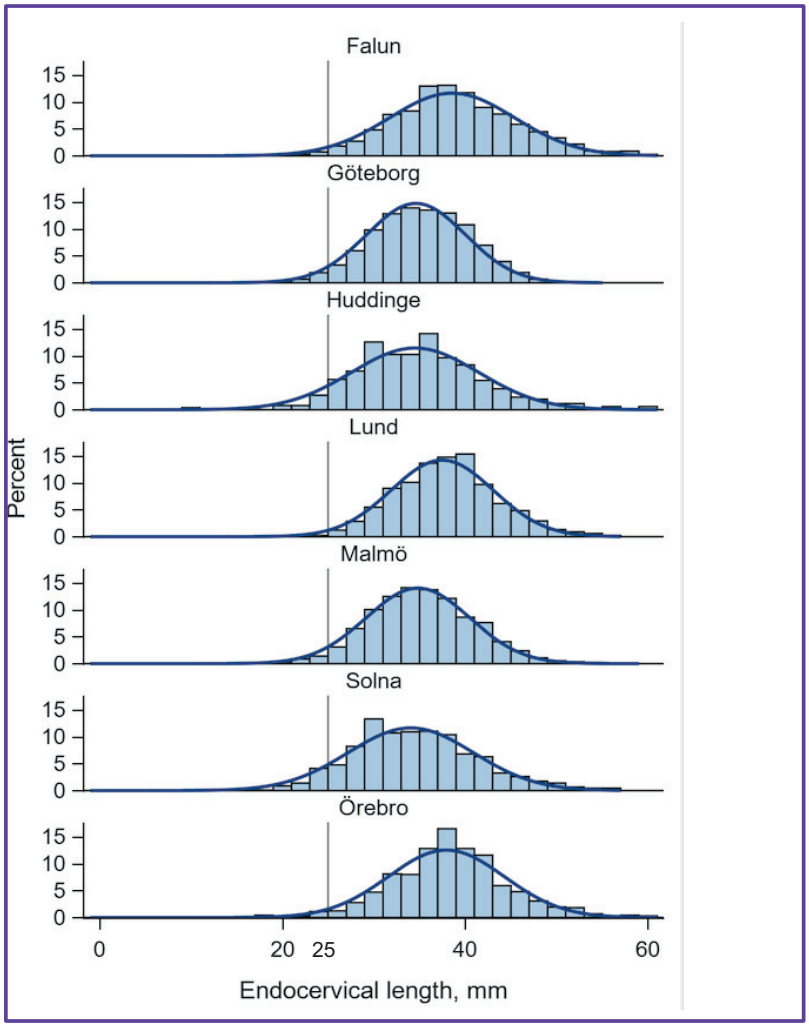


Figure 21 Distribution of endocervical length in the CERVIX study, per center

5.6 Screening or not

Today, there are screening programmes for women for the detection of breast cancer and cervical cancer in Sweden. Universal screening has also been discussed for colorectal cancer[193]. Screening is targeted at women within certain age ranges and is repeated at certain intervals. The number needed to screen (NNS) to prevent one death caused by breast cancer was 543 for women between 50 to 74 years and 3125 for women 40 to 49 years of age in the United States[194]. The corresponding figure for NNS in Sweden was 1252 for women between 40 and 49 years of age[195]. The NNS to prevent one death from cervical cancer screening was 1140[194].

Universal screening of cervical length measurement using transvaginal ultrasound and treatment with vaginal progesterone have been proposed for the prevention of PTD[63,196]. Several studies have shown that universal screening for cervical length is also a cost-effective strategy compared to risk-based screening or no screening[197-200]. If universal screening of cervical length to prevent sPTD is introduced for asymptomatic women in the general population, several points of view must be taken into account according to the WHO's ten criteria for establishing a good screening test[201]. The screening program should respond to a recognized need, there should be a defined target population, there should be scientific evidence of screening programme effectiveness, the program should integrate education, testing, clinical services and programme management, there should be quality assurance with mechanisms to minimize potential risks of screening, programme evaluation should be planned from the outset and the overall benefits of screening should outweigh the harm [201]. A few of these points will be discussed with regard to universal screening to assess cervical length for the prevention of sPTD: the scientific evidence of the effectiveness of the screening along with the benefits, as opposed to any harm, and the target population.

The evidence from our studies and other studies to screen the entire population with a low risk of PTD showed poor predictive values for PTD[55,66,68,69], (Paper I, Paper I). The ability of cervical length measurement to predict sPTD depends on which GW is chosen as an outcome variable. It can be discussed whether it is more important to prevent a PTD at <37 GW or <33 GW, or even earlier, where the neonatal benefits are larger. A reasonable NNS for cervical length measurement to identify one woman at risk of PTD, could be around 500. A false positive test result may initiate unnecessary treatments and anxiety and may cause extra follow-ups and sick leaves. How many false positive test results are acceptable to find one true positive? From

our point of view, the acceptable NNH should be around 30 to 35. In our studies most of the cervical length cut-offs did not fulfil these criteria. However, in the CERVIX study when the mathematically best cut-off (27 mm) was used at Cx2, the detection rate of sPTD <33 GW was 54%, NNH 35 and NNS 449, thus fulfilling the criteria above.

Screening of cervical length at 21+0 -23+6 GW (Cx2) and using the mathematically best cut-off point (<27 mm) identified 8% (510/6288) of the screened population as being at risk, while screening at 18+0 – 20+6 GW (Cx1) using the mathematically best cut-off (<29 mm) identified 15% (1638/11,072) as being at risk. Screening at Cx2 therefore seems preferable to screening at Cx1. However, screening at Cx2 will miss some women with a short cervical length at Cx1 who are at risk of a late miscarriage. Furthermore, screening at Cx2 will also lead to an extra visit for cervical length assessment since routine fetal anatomy screening is performed at around 18 GW. One possibility could be to assess the cervical length at the time of the routine fetal scan between 18 and 20+6 GW and to make a re-assessment of cervical length at between 21 and 23+6 GW and then offer treatment to those confirmed as being at risk.

This thesis has created a scientific basis for further discussions and studies as to whether universal cervical length screening or risk-based screening in Sweden should be implemented. There are still challenges before screening can be adopted. The current evidence of the effect of prophylactic treatment with progesterone (or other treatments) for women at risk of PTD is still under discussion. Furthermore, a health economic analysis on different screening strategies, universal or risk-based, is needed. This analysis may use the population data from the CERVIX study.

6. Conclusion

- A short cervical length measured in the second trimester is a risk factor for sPTD. The shorter the cervical length, the higher the risk of sPTD (Paper I, II)

- Cervical length in the second trimester provided no or poor discrimination between women who did and did not deliver preterm. However, it discriminated between women who did and did not deliver spontaneously preterm, the discriminative ability decreasing with advancing gestational age at delivery (Paper II)

- The ability of endocervical length to predict sPTD (Paper II) was better when cervical length was measured at 21+0-23+6 GW than at 18+0-20+6 GW and was better in women with no isthmus than in women with isthmus (Paper II)

- A change of cervical length between Cx1 and Cx2 did not discriminate better than cervical length at Cx1 or Cx2 alone (Paper II)

- The prevalence of short cervical length ≤25 mm was low in Paper I (0.5%) and considerably higher in Paper II (4.0%) (the CERVIX study)

- The matematically best cut-off of endocervical length was 37 mm to predict sPTD <34 GW (Paper I), 29 mm at Cx1 and 27 mm at Cx2 to predict sPTD <33 GW (Paper II). These cut-offs had sensitivity of 59%, 43% and 54% and AUC 0.69, 0.68 and 0.76, respectively

- Isthmus was present in 23% of cervical measurements at Cx1 and in 9% at Cx2 (Paper II)

- Inter-observer agreement and reliability of second trimester cervical length measurements differed substantially between examiner pairs. There were also varying levels of intra-observer measurement error, repeatability and reliability between the raters (Paper III)

- Inter-observer agreement for endocervical length was more accurate when the isthmus was judged to be absent (Paper III)

- Agreement that isthmus was absent was superior to agreement that isthmus was present (Paper II). Agreement that cervical length was > 25 mm was superior to agreement that cervical length was ≤25 mm (Paper III)

- Before universal cervical screening and treatment for the prevention of sPTD can be implemented in the general population in Sweden, an analysis of cost-effectiveness is needed. Ideally, an RCT on progesterone treatment in women with a short cervical length and with the same background characteristics (ethnicity, prior PTD etc) as the general Swedish population should be performed

7. Future perspective

This thesis showed that measurement of cervical length in the second trimester had a poor ability to discriminate between women who did or did not deliver preterm but discriminated better between women who did or did not have sPTD. Whether the discriminative ability is better in a high-risk group of women with a history of sPTD/PTD, previous late miscarriage, cervical incompetence or a history of cervical conization needs to be explored.

Because sPTD is associated with multiple mechanisms and etiologies, prevention in one case may not help in another case. A short cervix is a result of a process, which can start already in early pregnancy. If it is possible to find a marker to use together with the cervical measurement (or alone) it could give a new tool to find women at risk.

If universal screening is to be recommended, a health economic evaluation is needed with an analysis of screening costs, treatment costs, cost of sick leave and neonatal care etc. Furthermore, the low participation rate must be analysed and improved. Few studies have evaluated the woman's experience and their levels of acceptability of cervical measurement using transvaginal ultrasound. Such a study may include a study of the pain, discomfort and embarrassment caused by the scan.

To reduce the rate of PTD is a global challenge. The WHO's approaches to reducing PTD are: family planning (birth spacing, adolescent-friendly service); education and nutrition; prevention and screening for sexually transmitted infections; screening and management of high blood pressure and diabetes; behaviour change for lifestyle risks**; targeted care of women at increased risk for PTD**; smoking cessation; education to promote appropriate induction and caesarean section (Born too soon: www.who.int/maternal_child_adolescent/documents/born_too_soon/en/). In Sweden, although there is a low rate of PTD, we can do more to identify the women at increased risk of PTD and target care more precisely by establishing national guidelines for the prevention of this condition.

Acknowledgements

I am grateful to all the women who participated to these studies and made this work possible. Without facilities from University of Gothenburg, Sahlgrenska Academy, Departments of Gynecology and Obstetrics in Sahlgrenska University Hospital, Södra Älvsborg Hospital and other participating centers this thesis had never being feasible. The support from colleagues and friends has been invaluable for me, only some of whom are named below. I would like to give my special thanks to:

Ulla-Britt Wennerholm. My main supervisor who has taught me everything on how to conduct clinical trials. Your capacity, patience and endurance are amazing and I never stop honoring you. Thank you for all feedback and encouragement I have got during these years, I try to learn your positive attitude to solve problems and not only to see obstacles on the way.
Lil Valentin. My co-supervisor who has with a great generosity given her time and engagement to study details from the beginning to the end with enormous capacity and precision. I admire your talent in analytic thinking to discover the most important results among hundreds of analyses. I am grateful to have got this possibility to do these studies together.
Bo Jacobsson. My co-supervisor who has broad knowledge and vision in research about preterm delivery and gives perspectives to see my subject in relation to other studies. You have been supporting to find out solutions whenever needed.
Henrik Hagberg. My co-supervisor who always spreads intelligent and positive atmosphere in our meetings and gives honest feedback without delay when I ask for it.

Peter Lindgren, Helena Fadl, Jan Wesström, Carina Bejlum, Elisabeth Almström, Mona Söderlund and Lars Ladfors. My co-authors who have kindly helped with the study planning, recruiting and interpretations. I admire your ability to combine research with clinical reality.
All the ultrasonographers who have recruited the women at routine fetal scan and those midwifes who with enthusiasm have learnt to measure the cervical length and contributed to all data on it to this thesis. It would never have been possible without you! Specially thanks to *Martina, Christina, Anna, Sofia and Ammi* for always answering my inquiries and helping to control and pick up data.

Mattias Molin, Andreas Pehrsson and Nils Gunnar Pehrsson. Your understanding in numbers and relationships has been invaluable in analysing the study material and your pedagogic ability to explain the results to me.
Maria Gyhagen and Sigge Åkervall. Maria, your experience and enthusiasm in research infected and carried me through this process. I am grateful for your friendship and all inspiring discussions and wise advice you have given to me. Sigge, your sharing constructive criticism and interest in classical music with me.
Serney Bööj. Making it possible for me to combine the PhD studies with clinical work in your chief position at the clinic. It has been important to have your trust and understanding. And *Erica Cedervret Nilsson* continuing in the same manner.
FoU research unit of SÄS, at South Älvsborg Hospital, for creating a supporting and inspiring environment for clinical research. Special thanks to *Marie, Anneli, and Isabell.*
Annette Nattland, for professional help on the layout, sacrificing weekends with me for this. And *Gwyneth Olofsson* for checking the English of the main part of this dissertation; any mistakes in these acknowledgements are entirely mine. And *Eva Hessman* for professional help with the literature search.
Agneta Cederfors-Blohm for solving with practical issues and organizing around the studies, you are always so helpful and positive, nothing is impossible for you. And *Anja Andersson* for assistance with formalities and always smiling and encouraging me. *Malin Olsson* for solving administrative details around my research in SÄS.
All former and present colleagues who have helped me to solve the most difficult clinical problems together and standing by when I was doing my thesis.

Karin, Eva, Adalbjörg, Annika and Anna. My friends and former colleagues, with whom I can talk about everything between heaven and earth, life and death, work and family. And getting never-ending support, joy and hugs.
Carin, Kerstin, Anita and Elisabeth. You have been my first friends in this country in twenty years and coached me to through the language and cultural barriers. Your friendship has made me feel at home even here when you have generously given me a place in yours lives.
Katarina, Pernilla and Leticia. Our regular meetings have not only lowered the threshold to speak English but also made us friends who care about each other.
Sirkku, Anna and Tiina. We have shared the time studying medicine and family life together and now almost after thirty years and the Baltic see between us, the distance means nothing.
Marja and Tuula. Musicians and my friends who I have enjoyed not only in wonderful GSO concerts but also when we meet and can talk Finnish, it is so much

easier to express myself with you. I am happy that I have learned to know you, thanks to Bo and *Paulina*.
Per, Peter, Elisabet and Robin, Christina and Bengt, Marianne and Jan, Rosy and Jyrki, Jan and Sara, Laurette and Fransjohan. You have been models of wise persons for me, I trust and admire all of you.
The extended family in *Toiska*, where I have spent my summers since I was newborn. Living in that extraordinary community has taught me life in harmony with nature, cooperation, compromising skills, and accepting people as they are. Not to mention the relaxing sauna every night.
The worldwide knitting community in social media and in real life for inspiring me to create thinks with my hands and being the meditative break from all other duties. #lopapeysa, #kammebornia, #fruityknitting, #mariewallin, #jojilocat, to name some.

Virva, Mirka, and Pinja my sisters, and *Mikko* my brother. For sharing your lives with me: I am so happy to have you and your families by my side. Virva, our Sunday psychotherapy phone calls have been vital for me in the last 20 years back.
Reetta, my mother, who never stops encouraging all us five children. And who joins us together with our families during the holidays year after year. *Simo*, my father, who woke my interest in natural sciences and my need to explain what is behind everything.

My lovely sons and their dearest ones: *Eemu, Pauliina and Neeme, Uula and Linda, Aamos and Maja, Luukas*. What had I been without you? I am so proud of all of you!
Aarne, my love since high school and husband since 1987. Sharing scientific discussions with you has always been a part of my life, but now my increased understanding of research has joined us even closer to each other. Thank you for your love and encouragement the moments I almost lost my strength. You always stand by me.

This thesis was supported by grants from The Swedish Research Council (Dnr 2014-06998), Forskning och Utbildning (FoU) Södra Älvsborg and Research and Education Council (FoU-rådet) (VGFOUSA-864191, VGFOUSA-902741, VGFOUSA-902421, VGFOUSA-802751, VGFOUSA-611831).

Sammanfattning på svenska

Förtidsbörd definieras som en förlossning som sker före 37 fullgångna graviditetsveckor. Den utgör globalt en av de största utmaningarna inom förlossningsvården. I Sverige utgjorde förtidsbörd 5–6 % av förlossningarna mellan år 1986-2016. I många länder är frekvensen högre, t.ex. i USA 9,9% år 2017. Sjuklighet och dödlighet under nyföddhetsperioden är högre vid förtidsbörd än vid förlossning i fullgången tid. För tidigt födda barn löper också större risk att få bestående handikapp, såsom cerebral pares, jämfört med fullgångna barn. Ju tidigare förlossningen sker, desto större är riskerna för barnet. Många riskfaktorer för förtidsbörd har identifierats, men varannan kvinna med förtidsbörd har inga kända riskfaktorer. Kort livmodershals (cervix) är en riskfaktor associerad med spontan förtidsbörd. En del av studierna har visat att behandling med progesteron förlänger graviditeten och minskar den neonatala sjukligheten.

Syftet med denna avhandling var att studera sambandet mellan cervixlängden, mätt med vaginalt ultraljud, och förtidsbörd hos svenska kvinnor. Syftet var också att studera mätmetodens reproducerbarhet och pålitlighet, d.v.s. utfallet mellan olika undersökare och hos samma undersökare vid upprepade mätningar.

Två observationsstudier genomfördes där cervixlängden mättes med vaginalt ultraljud på kvinnor med enkelbörd i samband med rutinultraljud under den andra trimestern. Mätningarna utfördes av barnmorskor som efter teoretisk undervisning och praktisk upplärning certifierats för att utföra ultraljudsmätningarna av cervixlängden enligt fastställda kriterier för den aktuella studien. Den första studien (Paper I) genomfördes under 2012 och 2013 vid två kliniker (Göteborg och Trollhättan), och omfattade 2122 kvinnor. En mätning utfördes mellan graviditetsvecka 16 och 23. Kvinnor med kort cervix (≤25mm) rekryterades till en behandlingsstudie med progesteron. Den andra studien (CERVIX-studien, Paper II) genomfördes under 2014 till 2017 vid sju kliniker, och inkluderade 11 456 kvinnor. Cervix mättes första gången under graviditetsvecka 18 till 20 och därefter, med minst två veckors mellanrum, under graviditetsvecka 21 till 23. I CERVIX-studien var mätresultatet okänt för kvinnan och vårdgivaren.

Reproducerbarhetsstudien (Paper III) var en del av CERVIX-studien och bestod av två delstudier, LIVE-studien och CLIPS-studien. I LIVE-studien deltog sju par

certifierade barnmorskor, ett par från varje klinik. Varje barnmorskepar undersökte samma grupp av 24 till 30 kvinnor efter varandra utan att se varandras resultat. I CLIPS-studien deltog 16 barnmorskor som fick utföra mätningar på 93 videoclips från CERVIX-studien. Efter två månader upprepades proceduren på samma videoclips men i annan ordningsföljd.

I den första studien (Paper I) var medianvärdet för cervixlängden 39,0 mm. Kort cervix (≤25mm) återfanns hos 0,5 % av kvinnorna. Cervixlängden var korrelerad till spontan förtidsbörd; ju kortare cervix, desto större risk för förtidsbörd. I CERVIX-studien var medianvärdet 36,0 mm. Vid den första mätningen (graviditetsvecka 18 till 20) hade 4,0 % av kvinnorna en kort cervix (≤25 mm) och vid den andra (vecka 21 till 23) hade 4,4 % en kort cervix. I CERVIX studien framkom också att cervixlängden korrelerade till spontan förtidsbörd och korrelationen var starkare vid en tidig förlossning (före vecka 33) än vid förlossning i vecka 34 till 37. Prediktionen av förtidsbörd var bättre när mätningen gjordes under vecka 21 till 23, jämfört med mätning vecka 18 till 20. Bästa brytpunkt för att förutsäga spontan förtidsbörd före vecka 33 var en cervixlängd ≤27 mm, mätt under vecka 21 till 23. Med cervixlängden ≤27 mm, uppmätt under vecka 21 till 23 (8% av alla undersökta), kunde man identifiera 54 % (14/26) av de kvinnor som fick spontan förtidsbörd före vecka 33. För att hitta ett korrekt positivt testresultat (cervix ≤27 mm och föder vid <33 veckor) får 35 kvinnor falskt positivt testresultat (cervix ≤27 mm men föder vid ≥33 veckor) och totalt 449 kvinnor behöver undersökas (screenas).

För det bästa paret i LIVE-studien var medelskillnaden mellan undersökarnas mätningar 0,33 mm, och i 95 % av de parvisa undersökningarna var skillnaden mellan mätningarna ±5 mm. För det ”sämsta” paret var motsvarande siffror 0,7 mm och ±12 mm. I CLIPS-studien föll 95 % av skillnaderna mellan den bästa undersökarens två mätningar under 3,9 mm (”repeatability coefficient”) medan 95 % föll under 9,6 mm för den ”sämsta”. Medianvärdet för 16 undersökares ”repeatability coefficient” var 5,9 mm.

Slutsats: Kort cervixlängd är en riskfaktor för spontan förtidsbörd: ju kortare cervix, desto större risk. Som screeningmetod hade mätning av cervixlängd en måttlig förmåga att identifiera kvinnor som kom att föda för tidigt. Förmågan att förutsäga förtidsbörd var bättre vid mätning i sena graviditetsveckor (21 till 23 veckor) än i tidiga veckor (18 till 20). Trots undervisning, träning och certifiering varierade graden av överensstämmelse avsevärt mellan olika undersökare och för samma undersökare vid upprepade mätningar.

References

1. 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 obstetricia et gynecologica Scandinavica* 1977; **56**(3): 247-53.
2. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012; **379**(9832): 2162-72.
3. Blencowe H, Cousens S, Chou D, et al. Born too soon: the global epidemiology of 15 million preterm births. *Reproductive health* 2013; **10 Suppl 1**: S2.
4. Morken NH, Kallen K, Hagberg H, Jacobsson B. Preterm birth in Sweden 1973-2001: rate, subgroups, and effect of changing patterns in multiple births, maternal age, and smoking. *Acta obstetricia et gynecologica Scandinavica* 2005; **84**(6): 558-65.
5. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet* 2008; **371**(9606): 75-84.
6. Ananth CV, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2006; **19**(12): 773-82.
7. Romero R, Dey SK, Fisher SJ. Preterm labor: one syndrome, many causes. *Science (New York, NY)* 2014; **345**(6198): 760-5.
8. Esplin MS. Overview of spontaneous preterm birth: a complex and multifactorial phenotype. *Clinical obstetrics and gynecology* 2014; **57**(3): 518-30.
9. Ultra-ARG S. Riktlinjer för fetometri. 2010. https://www.sfog.se/start/arg/ultraljudsdiagnostik-ultra-arg/dokument/rekommendationer/ (accessed March 30 2019).
10. Saltvedt S, Almstrom H, Kublickas M, Reilly M, Valentin L, Grunewald C. Ultrasound dating at 12-14 or 15-20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of in-vitro fertilized pregnancies randomized to early or late dating scan. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2004; **24**(1): 42-50.
11. Selbing A, Kjessler B. Conceptual dating by ultrasonic measurement of the fetal biparietal diameter in early pregnancy. *Acta obstetricia et gynecologica Scandinavica* 1985; **64**(7): 593-7.
12. Sladkevicius P, Saltvedt S, Almstrom H, Kublickas M, Grunewald C, Valentin L. Ultrasound dating at 12-14 weeks of gestation. A prospective cross-validation of established dating formulae in in-vitro fertilized pregnancies. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2005; **26**(5): 504-11.
13. Fellman V, Hellstrom-Westas L, Norman M, et al. One-year survival of extremely preterm infants after active perinatal care in Sweden. *JAMA : the journal of the American Medical Association* 2009; **301**(21): 2225-33.
14. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* 2016; **388**(10063): 3027-35.

15. Crump C, Sundquist K, Sundquist J, Winkleby MA. Gestational age at birth and mortality in young adulthood. *JAMA : the journal of the American Medical Association* 2011; **306**(11): 1233-40.
16. Black RE, Levin C, Walker N, Chou D, Liu L, Temmerman M. Reproductive, maternal, newborn, and child health: key messages from Disease Control Priorities 3rd Edition. *Lancet* 2016; **388**(10061): 2811-24.
17. WHO. WHO recommendations on interventions to improve preterm birth outcomes. Geneva: World Health Organization. 2019-03-30 2015. http://apps.who.int/iris/bitstream/10665/183037/1/9789241508988_eng.pdf (accessed March 30 2019).
18. Norman M, Hallberg B, Abrahamsson T, et al. Association Between Year of Birth and 1-Year Survival Among Extremely Preterm Infants in Sweden During 2004-2007 and 2014-2016. *JAMA : the journal of the American Medical Association* 2019; **321**(12): 1188-99.
19. Ward RM, Beachy JC. Neonatal complications following preterm birth. *BJOG : an international journal of obstetrics and gynaecology* 2003; **110 Suppl 20**: 8-16.
20. Johanzon M, Odesjo H, Jacobsson B, Sandberg K, Wennerholm UB. Extreme preterm birth: onset of delivery and its effect on infant survival and morbidity. *Obstetrics and gynecology* 2008; **111**(1): 42-50.
21. Fritz T, Kallen K, Marsal K, Jacobsson B. Outcome of extremely preterm infants after iatrogenic or spontaneous birth. *Acta obstetricia et gynecologica Scandinavica* 2018; **97**(11): 1388-95.
22. Himmelmann K, Uvebrant P. The panorama of cerebral palsy in Sweden part XII shows that patterns changed in the birth years 2007-2010. *Acta paediatrica (Oslo, Norway : 1992)* 2018; **107**(3): 462-8.
23. Serenius F, Kallen K, Blennow M, et al. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. *JAMA : the journal of the American Medical Association* 2013; **309**(17): 1810-20.
24. Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *The New England journal of medicine* 2008; **359**(3): 262-73.
25. Iams JD, Goldenberg RL, Mercer BM, et al. The preterm prediction study: can low-risk women destined for spontaneous preterm birth be identified? *American journal of obstetrics and gynecology* 2001; **184**(4): 652-5.
26. Copper RL, Goldenberg RL, Das A, et al. The preterm prediction study: maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *American journal of obstetrics and gynecology* 1996; **175**(5): 1286-92.
27. Ehrenberg HM, Dierker L, Milluzzi C, Mercer BM. Low maternal weight, failure to thrive in pregnancy, and adverse pregnancy outcomes. *American journal of obstetrics and gynecology* 2003; **189**(6): 1726-30.
28. Esplin MS, O'Brien E, Fraser A, et al. Estimating recurrence of spontaneous preterm delivery. *Obstetrics and gynecology* 2008; **112**(3): 516-23.
29. Hendler I, Goldenberg RL, Mercer BM, et al. The Preterm Prediction Study: association between maternal body mass index and spontaneous and indicated preterm birth. *American journal of obstetrics and gynecology* 2005; **192**(3): 882-6.
30. Iams JD, Romero R., Creasy, R.K. Creasy & Resnik's Maternal- Fetal Medicine: Priniciple and Practice. In: Creasy R. K. RR, Iams J.D, ed. Creasy & Resnik's Maternal-Fetal Medicine, Priniciple and Practice. 6th ed: Saunders; 2009: 546-82.
31. Koullali B, Oudijk MA, Nijman TA, Mol BW, Pajkrt E. Risk assessment and management to prevent preterm birth. *Seminars in fetal & neonatal medicine* 2016; **21**(2): 80-8.

32. Laughon SK, Albert PS, Leishear K, Mendola P. The NICHD Consecutive Pregnancies Study: recurrent preterm delivery by subtype. *American journal of obstetrics and gynecology* 2014; **210**(2): 131.e1-8.
33. Meis PJ, Goldenberg RL, Mercer BM, et al. The preterm prediction study: risk factors for indicated preterm births. Maternal-Fetal Medicine Units Network of the National Institute of Child Health and Human Development. *American journal of obstetrics and gynecology* 1998; **178**(3): 562-7.
34. Mercer BM, Goldenberg RL, Moawad AH, et al. The preterm prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *American journal of obstetrics and gynecology* 1999; **181**(5 Pt 1): 1216-21.
35. Oliver-Williams C, Fleming M, Wood AM, Smith G. Previous miscarriage and the subsequent risk of preterm birth in Scotland, 1980-2008: a historical cohort study. *BJOG : an international journal of obstetrics and gynaecology* 2015; **122**(11): 1525-34.
36. Phillips C, Velji Z, Hanly C, Metcalfe A. Risk of recurrent spontaneous preterm birth: a systematic review and meta-analysis. *BMJ open* 2017; **7**(6): e015402.
37. Simonsen SE, Lyon JL, Stanford JB, Porucznik CA, Esplin MS, Varner MW. Risk factors for recurrent preterm birth in multiparous Utah women: a historical cohort study. *BJOG : an international journal of obstetrics and gynaecology* 2013; **120**(7): 863-72.
38. Swingle HM, Colaizy TT, Zimmerman MB, Morriss FH, Jr. Abortion and the risk of subsequent preterm birth: a systematic review with meta-analyses. *The Journal of reproductive medicine* 2009; **54**(2): 95-108.
39. Simhan HN, Iams J.D., Romero R. Preterm Birth. In: E.G. G, ed. Obstetrics: Normal and problem pregnancies. 6th ed: Elsevier; 2012: 627-56.
40. Adams Waldorf KM, Singh N, Mohan AR, et al. Uterine overdistention induces preterm labor mediated by inflammation: observations in pregnant women and nonhuman primates. *American journal of obstetrics and gynecology* 2015; **213**(6): 830.e1-.e19.
41. Cnattingius S, Villamor E, Johansson S, et al. Maternal obesity and risk of preterm delivery. *JAMA : the journal of the American Medical Association* 2013; **309**(22): 2362-70.
42. Kazemier BM, Buijs PE, Mignini L, Limpens J, de Groot CJ, Mol BW. Impact of obstetric history on the risk of spontaneous preterm birth in singleton and multiple pregnancies: a systematic review. *BJOG : an international journal of obstetrics and gynaecology* 2014; **121**(10): 1197-208; discussion 209.
43. Jakobsson M, Gissler M, Paavonen J, Tapper AM. Loop electrosurgical excision procedure and the risk for preterm birth. *Obstetrics and gynecology* 2009; **114**(3): 504-10.
44. Weinmann S, Naleway A, Swamy G, et al. Pregnancy Outcomes after Treatment for Cervical Cancer Precursor Lesions: An Observational Study. *PloS one* 2017; **12**(1): e0165276.
45. Henderson JJ, McWilliam OA, Newnham JP, Pennell CE. Preterm birth aetiology 2004-2008. Maternal factors associated with three phenotypes: spontaneous preterm labour, preterm pre-labour rupture of membranes and medically indicated preterm birth. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2012; **25**(6): 642-7.
46. Romero R, Espinoza J, Kusanovic JP, et al. The preterm parturition syndrome. *BJOG : an international journal of obstetrics and gynaecology* 2006; **113 Suppl 3**: 17-42.
47. Gray's Anatomy. 37th ed: Churchill Livingstone; 1989: 1440-44

48. Ross MH, Reith E.J, Romrell L.J. Histology, A Text and Atlas. International edition, 2nd ed: Williams & Wilkins; 1989.
49. Read CP, Word RA, Ruscheinsky MA, Timmons BC, Mahendroo MS. Cervical remodeling during pregnancy and parturition: molecular characterization of the softening phase in mice. *Reproduction (Cambridge, England)* 2007; **134**(2): 327-40.
50. Timmons B, Akins M, Mahendroo M. Cervical remodeling during pregnancy and parturition. *Trends in endocrinology and metabolism: TEM* 2010; **21**(6): 353-61.
51. Osmers R, Rath W, Pflanz MA, Kuhn W, Stuhlsatz HW, Szeverenyi M. Glycosaminoglycans in cervical connective tissue during pregnancy and parturition. *Obstetrics and gynecology* 1993; **81**(1): 88-92.
52. Goldman S, Weiss A, Eyali V, Shalev E. Differential activity of the gelatinases (matrix metalloproteinases 2 and 9) in the fetal membranes and decidua, associated with labour. *Molecular human reproduction* 2003; **9**(6): 367-73.
53. Bergelin I, Valentin L. Patterns of normal change in cervical length and width during pregnancy in nulliparous women: a prospective, longitudinal ultrasound study. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2001; **18**(3): 217-22.
54. Souka AP, Papastefanou I, Michalitsi V, et al. Cervical length changes from the first to second trimester of pregnancy, and prediction of preterm birth by first-trimester sonographic cervical measurement. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2011; **30**(7): 997-1002.
55. Wulff CB, Rode L, Rosthoj S, Hoseth E, Petersen OB, Tabor A. Transvaginal sonographic cervical length in first and second trimesters in a low-risk population: a prospective study. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2018; **51**(5): 604-13.
56. Zorzoli A, Soliani A, Perra M, Caravelli E, Galimberti A, Nicolini U. Cervical changes throughout pregnancy as assessed by transvaginal sonography. *Obstetrics and gynecology* 1994; **84**(6): 960-4.
57. Gramellini D, Fieni S, Molina E, Berretta R, Vadora E. Transvaginal sonographic cervical length changes during normal pregnancy. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2002; **21**(3): 227-32; quiz 34-5.
58. Albayrak M, Ozdemir I, Koc O, Coskun E. Can maternal height predict shorter cervical length in asymptomatic low-risk pregnant women? *European journal of obstetrics, gynecology, and reproductive biology* 2011; **157**(2): 161-5.
59. van der Ven AJ, van Os MA, Kleinrouweler CE, et al. Is cervical length associated with maternal characteristics? *European journal of obstetrics, gynecology, and reproductive biology* 2015; **188**: 12-6.
60. Heath VC, Southall TR, Souka AP, Novakov A, Nicolaides KH. Cervical length at 23 weeks of gestation: relation to demographic characteristics and previous obstetric history. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 1998; **12**(5): 304-11.
61. Palma-Dias RS, Fonseca MM, Stein NR, Schmidt AP, Magalhaes JA. Relation of cervical length at 22-24 weeks of gestation to demographic characteristics and obstetric history. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas* 2004; **37**(5): 737-44.
62. Drakeley AJ, Roberts D, Alfirevic Z. Cervical stitch (cerclage) for preventing pregnancy loss in women. *The Cochrane database of systematic reviews* 2003; (1): Cd003253.

63. Campbell S. Prevention of spontaneous preterm birth: universal cervical length assessment and vaginal progesterone in women with a short cervix: time for action! *American journal of obstetrics and gynecology* 2018; **218**(2): 151-8.
64. Iams JD, Goldenberg RL, Meis PJ, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *The New England journal of medicine* 1996; **334**(9): 567-72.
65. van der Ven J, van Os MA, Kazemier BM, et al. The capacity of mid-pregnancy cervical length to predict preterm birth in low-risk women: a national cohort study. *Acta obstetricia et gynecologica Scandinavica* 2015; **94**(11): 1223-34.
66. Esplin MS, Elovitz MA, Iams JD, et al. Predictive Accuracy of Serial Transvaginal Cervical Lengths and Quantitative Vaginal Fetal Fibronectin Levels for Spontaneous Preterm Birth Among Nulliparous Women. *JAMA : the journal of the American Medical Association* 2017; **317**(10): 1047-56.
67. Heath VC, Southall TR, Souka AP, Elisseou A, Nicolaides KH. Cervical length at 23 weeks of gestation: prediction of spontaneous preterm delivery. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 1998; **12**(5): 312-7.
68. Leung TN, Pang MW, Leung TY, Poon CF, Wong SM, Lau TK. Cervical length at 18-22 weeks of gestation for prediction of spontaneous preterm delivery in Hong Kong Chinese women. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2005; **26**(7): 713-7.
69. Taipale P, Hiilesmaa V. Sonographic measurement of uterine cervix at 18-22 weeks' gestation and the risk of preterm delivery. *Obstetrics and gynecology* 1998; **92**(6): 902-7.
70. Son M, Miller ES. Predicting preterm birth: Cervical length and fetal fibronectin. *Seminars in perinatology* 2017; **41**(8): 445-51.
71. Domin CM, Smith EJ, Terplan M. Transvaginal ultrasonographic measurement of cervical length as a predictor of preterm birth: a systematic review with meta-analysis. *Ultrasound quarterly* 2010; **26**(4): 241-8.
72. Goldenberg RL, Thom E, Moawad AH, Johnson F, Roberts J, Caritis SN. The preterm prediction study: fetal fibronectin, bacterial vaginosis, and peripartum infection. NICHD Maternal Fetal Medicine Units Network. *Obstetrics and gynecology* 1996; **87**(5 Pt 1): 656-60.
73. Peaceman AM, Andrews WW, Thorp JM, et al. Fetal fibronectin as a predictor of preterm birth in patients with symptoms: a multicenter trial. *American journal of obstetrics and gynecology* 1997; **177**(1): 13-8.
74. Abbott DS, Radford SK, Seed PT, Tribe RM, Shennan AH. Evaluation of a quantitative fetal fibronectin test for spontaneous preterm birth in symptomatic women. *American journal of obstetrics and gynecology* 2013; **208**(2): 122.e1-6.
75. Abbott DS, Hezelgrave NL, Seed PT, et al. Quantitative fetal fibronectin to predict preterm birth in asymptomatic women at high risk. *Obstetrics and gynecology* 2015; **125**(5): 1168-76.
76. Carvalho MH, Bittar RE, Brizot ML, Maganha PP, Borges da Fonseca ES, Zugaib M. Cervical length at 11-14 weeks' and 22-24 weeks' gestation evaluated by transvaginal sonography, and gestational age at delivery. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2003; **21**(2): 135-9.
77. Kazemier BM, Miller ES, Grobman WA, Mol BW. Variation in preterm birth rate and the role of short cervical length across two populations: a comparative cohort study.

Journal of perinatology : official journal of the California Perinatal Association 2016; **36**(7): 516-21.
78. Guzman ER, Walters C, Ananth CV, et al. A comparison of sonographic cervical parameters in predicting spontaneous preterm birth in high-risk singleton gestations. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2001; **18**(3): 204-10.
79. Owen J, Yost N, Berghella V, et al. Mid-trimester endovaginal sonography in women at high risk for spontaneous preterm birth. *JAMA : the journal of the American Medical Association* 2001; **286**(11): 1340-8.
80. To MS, Fonseca EB, Molina FS, Cacho AM, Nicolaides KH. Maternal characteristics and cervical length in the prediction of spontaneous early preterm delivery in twins. *American journal of obstetrics and gynecology* 2006; **194**(5): 1360-5.
81. Iams JD, Romero R, Culhane JF, Goldenberg RL. Primary, secondary, and tertiary interventions to reduce the morbidity and mortality of preterm birth. *Lancet* 2008; **371**(9607): 164-75.
82. Matei A, Saccone G, Vogel JP, Armson AB. Primary and secondary prevention of preterm birth: a review of systematic reviews and ongoing randomized controlled trials. *European journal of obstetrics, gynecology, and reproductive biology* 2019.
83. Romero R, Gomez R, Ghezzi F, et al. A fetal systemic inflammatory response is followed by the spontaneous onset of preterm parturition. *American journal of obstetrics and gynecology* 1998; **179**(1): 186-93.
84. Askie LM, Duley L, Henderson-Smart DJ, Stewart LA. Antiplatelet agents for prevention of pre-eclampsia: a meta-analysis of individual patient data. *Lancet* 2007; **369**(9575): 1791-8.
85. Berghella V, Figueroa D, Szychowski JM, et al. 17-alpha-hydroxyprogesterone caproate for the prevention of preterm birth in women with prior preterm birth and a short cervical length. *American journal of obstetrics and gynecology* 2010; **202**(4): 351.e1-6.
86. Hassan SS, Romero R, Vidyadhari D, et al. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2011; **38**(1): 18-31.
87. Norman JE, Marlow N, Messow CM, et al. Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial. *Lancet* 2016; **387**(10033): 2106-16.
88. Romero R, Conde-Agudelo A, Da Fonseca E, et al. Vaginal progesterone for preventing preterm birth and adverse perinatal outcomes in singleton gestations with a short cervix: a meta-analysis of individual patient data. *American journal of obstetrics and gynecology* 2018; **218**(2): 161-80.
89. Berghella V, Rafael TJ, Szychowski JM, Rust OA, Owen J. Cerclage for short cervix on ultrasonography in women with singleton gestations and previous preterm birth: a meta-analysis. *Obstetrics and gynecology* 2011; **117**(3): 663-71.
90. Goya M, Pratcorona L, Merced C, et al. Cervical pessary in pregnant women with a short cervix (PECEP): an open-label randomised controlled trial. *Lancet* 2012; **379**(9828): 1800-6.
91. Nicolaides KH, Syngelaki A, Poon LC, et al. A Randomized Trial of a Cervical Pessary to Prevent Preterm Singleton Birth. *The New England journal of medicine* 2016; **374**(11): 1044-52.
92. Jarde A, Lutsiv O, Beyene J, McDonald SD. Vaginal progesterone, oral progesterone, 17-OHPC, cerclage, and pessary for preventing preterm birth in at-risk singleton

pregnancies: an updated systematic review and network meta-analysis. *BJOG : an international journal of obstetrics and gynaecology* 2019; **126**(5): 556-67.
93. Alfirevic Z, Stampalija T, Medley N. Cervical stitch (cerclage) for preventing preterm birth in singleton pregnancy. *The Cochrane database of systematic reviews* 2017; **6**: Cd008991.
94. Jorgensen AL, Alfirevic Z, Tudur Smith C, Williamson PR. Cervical stitch (cerclage) for preventing pregnancy loss: individual patient data meta-analysis. *BJOG : an international journal of obstetrics and gynaecology* 2007; **114**(12): 1460-76.
95. Hui SY, Chor CM, Lau TK, Lao TT, Leung TY. Cerclage pessary for preventing preterm birth in women with a singleton pregnancy and a short cervix at 20 to 24 weeks: a randomized controlled trial. *American journal of perinatology* 2013; **30**(4): 283-8.
96. Saccone G, Ciardulli A, Xodo S, et al. Cervical pessary for preventing preterm birth in twin pregnancies with short cervical length: a systematic review and meta-analysis. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2017; **30**(24): 2918-25.
97. Zheng L, Dong J, Dai Y, et al. Cervical pessaries for the prevention of preterm birth: a systematic review and meta-analysis. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2019; **32**(10): 1654-63.
98. Karbasian N, Sheikh M, Pirjani R, Hazrati S, Tara F, Hantoushzadeh S. Combined treatment with cervical pessary and vaginal progesterone for the prevention of preterm birth: A randomized clinical trial. *The journal of obstetrics and gynaecology research* 2016; **42**(12): 1673-9.
99. Cruz-Melguizo S, San-Frutos L, Martinez-Payo C, et al. Cervical Pessary Compared With Vaginal Progesterone for Preventing Early Preterm Birth: A Randomized Controlled Trial. *Obstetrics and gynecology* 2018; **132**(4): 907-15.
100. Meis PJ, Klebanoff M, Thom E, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. *The New England journal of medicine* 2003; **348**(24): 2379-85.
101. Grobman WA, Thom EA, Spong CY, et al. 17 alpha-hydroxyprogesterone caproate to prevent prematurity in nulliparas with cervical length less than 30 mm. *American journal of obstetrics and gynecology* 2012; **207**(5): 390.e1-8.
102. Cetingoz E, Cam C, Sakalli M, Karateke A, Celik C, Sancak A. Progesterone effects on preterm birth in high-risk pregnancies: a randomized placebo-controlled trial. *Archives of gynecology and obstetrics* 2011; **283**(3): 423-9.
103. Fonseca EB, Celik E, Parra M, Singh M, Nicolaides KH. Progesterone and the risk of preterm birth among women with a short cervix. *The New England journal of medicine* 2007; **357**(5): 462-9.
104. O'Brien JM, Adair CD, Lewis DF, et al. Progesterone vaginal gel for the reduction of recurrent preterm birth: primary results from a randomized, double-blind, placebo-controlled trial. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2007; **30**(5): 687-96.
105. van Os MA, van der Ven AJ, Kleinrouweler CE, et al. Preventing Preterm Birth with Progesterone in Women with a Short Cervical Length from a Low-Risk Population: A Multicenter Double-Blind Placebo-Controlled Randomized Trial. *American journal of perinatology* 2015; **32**(10): 993-1000.

106. Dodd JM, Grivell RM, CM OB, Dowswell T, Deussen AR. Prenatal administration of progestogens for preventing spontaneous preterm birth in women with a multiple pregnancy. *The Cochrane database of systematic reviews* 2017; **10**: Cd012024.
107. Berghella V, Ciardulli A, Rust OA, et al. Cerclage for sonographic short cervix in singleton gestations without prior spontaneous preterm birth: systematic review and meta-analysis of randomized controlled trials using individual patient-level data. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2017; **50**(5): 569-77.
108. Dodd JM, Jones L, Flenady V, Cincotta R, Crowther CA. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. *The Cochrane database of systematic reviews* 2013; (7): Cd004947.
109. Cho HJ, Roh HJ. Correlation Between Cervical Lengths Measured by Transabdominal and Transvaginal Sonography for Predicting Preterm Birth. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2016; **35**(3): 537-44.
110. Cicero S, Skentou C, Souka A, To MS, Nicolaides KH. Cervical length at 22-24 weeks of gestation: comparison of transvaginal and transperineal-translabial ultrasonography. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2001; **17**(4): 335-40.
111. Gauthier T, Marin B, Garuchet-Bigot A, et al. Transperineal versus transvaginal ultrasound cervical length measurement and preterm labor. *Archives of gynecology and obstetrics* 2014; **290**(3): 465-9.
112. Hernandez-Andrade E, Romero R, Ahn H, et al. Transabdominal evaluation of uterine cervical length during pregnancy fails to identify a substantial number of women with a short cervix. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2012; **25**(9): 1682-9.
113. Marren AJ, Mogra R, Pedersen LH, Walter M, Ogle RF, Hyett JA. Ultrasound assessment of cervical length at 18-21 weeks' gestation in an Australian obstetric population: comparison of transabdominal and transvaginal approaches. *The Australian & New Zealand journal of obstetrics & gynaecology* 2014; **54**(3): 250-5.
114. Roh HJ, Ji YI, Jung CH, Jeon GH, Chun S, Cho HJ. Comparison of cervical lengths using transabdominal and transvaginal sonography in midpregnancy. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2013; **32**(10): 1721-8.
115. Stone PR, Chan EH, McCowan LM, Taylor RS, Mitchell JM. Transabdominal scanning of the cervix at the 20-week morphology scan: comparison with transvaginal cervical measurements in a healthy nulliparous population. *The Australian & New Zealand journal of obstetrics & gynaecology* 2010; **50**(6): 523-7.
116. To MS, Skentou C, Cicero S, Nicolaides KH. Cervical assessment at the routine 23-weeks' scan: problems with transabdominal sonography. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2000; **15**(4): 292-6.
117. Yazici G, Yildiz A, Tiras MB, Arslan M, Kanik A, Oz U. Comparison of transperineal and transvaginal sonography in predicting preterm delivery. *Journal of clinical ultrasound : JCU* 2004; **32**(5): 225-30.
118. Berghella V, Tolosa JE, Kuhlman K, Weiner S, Bolognese RJ, Wapner RJ. Cervical ultrasonography compared with manual examination as a predictor of preterm delivery. *American journal of obstetrics and gynecology* 1997; **177**(4): 723-30.

119. Onderoglu LS. Digital examination and transperineal ultrasonographic measurement of cervical length to assess risk of preterm delivery. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 1997; **59**(3): 223-8.
120. Baxter JK, Adair CD, Paidas MJ, et al. Assessment of a cervicometer compared to transvaginal ultrasound in identifying women with a short cervical length: a multicenter study. *American journal of obstetrics and gynecology* 2016; **215**(2): 229.e1-7.
121. Preterm labour and birth. 2015. https://www.nice.org.uk/guidance/ng25/chapter/Recommendations#prophylactic-vaginal-progesterone-and-prophylactic-cervical-cerclage (accessed March 30 2019).
122. Lim K, Butt K, Crane JM. SOGC Clinical Practice Guideline. Ultrasonographic cervical length assessment in predicting preterm birth in singleton pregnancies. *Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC* 2011; **33**(5): 486-99.
123. SMFM (Society for Maternal-Fetal MEdicine Publications Committee waoVB. Progesterone and preterm birth prevention: translating clinical trials data into clinical practice. *American journal of obstetrics and gynecology* 2012; **206**(5): 376-86.
124. Stamilio D, Carlson LM. Transabdominal ultrasound is appropriate. *American journal of obstetrics and gynecology* 2016; **215**(6): 739-43.e1.
125. Hernandez-Andrade E, Romero R, Korzeniewski SJ, et al. Cervical strain determined by ultrasound elastography and its association with spontaneous preterm delivery. *Journal of perinatal medicine* 2014; **42**(2): 159-69.
126. Khalil MR, Thorsen P, Uldbjerg N. Cervical ultrasound elastography may hold potential to predict risk of preterm birth. *Danish medical journal* 2013; **60**(1): A4570.
127. Sabiani L, Haumonte JB, Loundou A, et al. Cervical HI-RTE elastography and pregnancy outcome: a prospective study. *European journal of obstetrics, gynecology, and reproductive biology* 2015; **186**: 80-4.
128. von Schoning D, Fischer T, von Tucher E, et al. Cervical sonoelastography for improving prediction of preterm birth compared with cervical length measurement and fetal fibronectin test. *Journal of perinatal medicine* 2015; **43**(5): 531-6.
129. Banos N, Murillo-Bravo C, Julia C, et al. Mid-trimester sonographic cervical consistency index to predict spontaneous preterm birth in a low-risk population. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2018; **51**(5): 629-36.
130. Parra-Saavedra M, Gomez L, Barrero A, Parra G, Vergara F, Navarro E. Prediction of preterm birth using the cervical consistency index. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2011; **38**(1): 44-51.
131. Barber MA, Medina M, Cabrera F, Romero A, Valle L, Garcia-Hernandez JA. Cervical length vs VOCAL cervical volume for predicting pre-term delivery in asymptomatic women at 20-22 weeks' pregnancy. *Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology* 2012; **32**(7): 648-51.
132. Bega G, Lev-Toaff A, Kuhlman K, et al. Three-dimensional multiplanar transvaginal ultrasound of the cervix in pregnancy. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2000; **16**(4): 351-8.
133. Severi FM, Bocchi C, Florio P, Picciolini E, D'Aniello G, Petraglia F. Comparison of two-dimensional and three-dimensional ultrasound in the assessment of the cervix to predict preterm delivery. *Ultrasound in medicine & biology* 2003; **29**(9): 1261-5.

134. Dziadosz M, Bennett TA, Dolin C, et al. Uterocervical angle: a novel ultrasound screening tool to predict spontaneous preterm birth. *American journal of obstetrics and gynecology* 2016; **215**(3): 376.e1-7.
135. Farras Llobet A, Regincos Marti L, Higueras T, et al. The uterocervical angle and its relationship with preterm birth. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2018; **31**(14): 1881-4.
136. Tekesin I, Wallwiener D, Schmidt S. The value of quantitative ultrasound tissue characterization of the cervix and rapid fetal fibronectin in predicting preterm delivery. 2005; **33**(5): 383-91.
137. Fukami T, Ishihara K, Sekiya T, Araki T. Is transvaginal ultrasonography at mid-trimester useful for predicting early spontaneous preterm birth? *Journal of Nippon Medical School = Nippon Ika Daigaku zasshi* 2003; **70**(2): 135-40.
138. Pires CR, Moron AF, Mattar R, Diniz AL, Andrade SG, Bussamra LC. Cervical gland area as an ultrasonographic marker for preterm delivery. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2006; **93**(3): 214-9.
139. De Diego R, Sabria J, Vela A, Rodriguez D, Gomez MD. Role of 3-dimensional power Doppler sonography in differentiating pregnant women with threatened preterm labor from those with an asymptomatic short cervix. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2014; **33**(4): 673-9.
140. Samutchaikij T, Pitukkijronnakorn S, Panburana P. Normal reference of cervical blood perfusion in pregnancy. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* 2014; **97**(4): 369-73.
141. Parra-Cordero M, Sepulveda-Martinez A, Rencoret G, Valdes E, Pedraza D, Munoz H. Is there a role for cervical assessment and uterine artery Doppler in the first trimester of pregnancy as a screening test for spontaneous preterm delivery? *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2014; **43**(3): 291-6.
142. Kagan KO, Sonek J. How to measure cervical length. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2015; **45**(3): 358-62.
143. Crane JM, Delaney T, Hutchens D. Transvaginal ultrasonography in the prediction of preterm birth after treatment for cervical intraepithelial neoplasia. *Obstetrics and gynecology* 2006; **107**(1): 37-44.
144. Haas DM, Parker CB, Wing DA, et al. A description of the methods of the Nulliparous Pregnancy Outcomes Study: monitoring mothers-to-be (nuMoM2b). *American journal of obstetrics and gynecology* 2015; **212**(4): 539.e1-.e24.
145. To MS, Skentou C, Liao AW, Cacho A, Nicolaides KH. Cervical length and funneling at 23 weeks of gestation in the prediction of spontaneous early preterm delivery. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2001; **18**(3): 200-3.
146. DeFranco EA, O'Brien JM, Adair CD, et al. Vaginal progesterone is associated with a decrease in risk for early preterm birth and improved neonatal outcome in women with a short cervix: a secondary analysis from a randomized, double-blind, placebo-controlled trial. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2007; **30**(5): 697-705.

147. Boelig RC, Feltovich H, Spitz JL, Toland G, Berghella V, Iams JD. Assessment of Transvaginal Ultrasound Cervical Length Image Quality. *Obstetrics and gynecology* 2017; **129**(3): 536-41.
148. van Os MA, van der Ven AJ, Bloemendaal PM, et al. Effect of e-learning on quality of cervical-length measurements. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2015; **46**(3): 327-31.
149. Chory MK, Schnettler WT, March M, Hacker MR, Modest AM, Rodriguez D. ACES: Accurate Cervical Evaluation With Sonography. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* 2016; **35**(1): 25-8.
150. Iams JD, Grobman WA, Lozitska A, et al. Adherence to criteria for transvaginal ultrasound imaging and measurement of cervical length. *American journal of obstetrics and gynecology* 2013; **209**(4): 365.e1-5.
151. Yost NP, Bloom SL, Twickler DM, Leveno KJ. Pitfalls in ultrasonic cervical length measurement for predicting preterm birth. *Obstetrics and gynecology* 1999; **93**(4): 510-6.
152. van Os MA, Kleinrouweler CE, Schuit E, et al. Influence of cut-off value on prevalence of short cervical length. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2017; **49**(3): 330-6.
153. Barros-Silva J, Pedrosa AC, Matias A. Sonographic measurement of cervical length as a predictor of preterm delivery: a systematic review. *Journal of perinatal medicine* 2014; **42**(3): 281-93.
154. Honest H, Bachmann LM, Coomarasamy A, Gupta JK, Kleijnen J, Khan KS. Accuracy of cervical transvaginal sonography in predicting preterm birth: a systematic review. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2003; **22**(3): 305-22.
155. Davies G, Ottenhof C, Woodman M, et al. Cervix length and relaxin as predictors of preterm birth. *Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC* 2008; **30**(12): 1124-31.
156. Matijevic R, Grgic O, Knezevic M. Vaginal pH versus cervical length in the mid-trimester as screening predictors of preterm labor in a low-risk population. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2010; **111**(1): 41-4.
157. Heath VC, Daskalakis G, Zagaliki A, Carvalho M, Nicolaides KH. Cervicovaginal fibronectin and cervical length at 23 weeks of gestation: relative risk of early preterm delivery. *BJOG : an international journal of obstetrics and gynaecology* 2000; **107**(10): 1276-81.
158. Jwala S, Tran TL, Terenna C, et al. Evaluation of additive effect of quantitative fetal fibronectin to cervical length for prediction of spontaneous preterm birth among asymptomatic low-risk women. *Acta obstetricia et gynecologica Scandinavica* 2016; **95**(8): 948-55.
159. Barber MA, Eguiluz I, Plasencia W, Medina M, Valle L, Garcia JA. Preterm delivery and ultrasound measurement of cervical length in Gran Canaria, Spain. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2010; **108**(1): 58-60.
160. Hebbar S, Samjhana K. Role of mid-trimester transvaginal cervical ultrasound in prediction of preterm delivery. *The Medical journal of Malaysia* 2006; **61**(3): 307-11.
161. Hibbard JU, Tart M, Moawad AH. Cervical length at 16-22 weeks' gestation and risk for preterm delivery. *Obstetrics and gynecology* 2000; **96**(6): 972-8.

162. Dilek TU, Yazici G, Gurbuz A, et al. Progressive cervical length changes versus single cervical length measurement by transvaginal ultrasound for prediction of preterm delivery. *Gynecologic and obstetric investigation* 2007; **64**(4): 175-9.
163. Kottner J, Audige L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of clinical epidemiology* 2011; **64**(1): 96-106.
164. Bland JM, Altman DG. Measurement error. *BMJ (Clinical research ed)* 1996; **312**(7047): 1654.
165. Franca C, Carraca T, Monteiro SB, et al. Inter- and intra-observer variability in cervical measurement by ultrasound in the first and second trimesters of pregnancy: does it matter? *Journal of perinatal medicine* 2015; **43**(1): 67-73.
166. Valentin L, Bergelin I. Intra- and interobserver reproducibility of ultrasound measurements of cervical length and width in the second and third trimesters of pregnancy. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2002; **20**(3): 256-62.
167. Stephansson O, Petersson K, Bjork C, Conner P, Wikstrom AK. The Swedish Pregnancy Register - for quality of care improvement and research. *Acta obstetricia et gynecologica Scandinavica* 2018; **97**(4): 466-76.
168. Petersson K, Persson M, Lindkvist M, et al. Internal validity of the Swedish Maternal Health Care Register. *BMC health services research* 2014; **14**: 364.
169. Cnattingius S, Ericson A, Gunnarskog J, Kallen B. A quality study of a medical birth registry. *Scandinavian journal of social medicine* 1990; **18**(2): 143-8.
170. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC public health* 2011; **11**: 450.
171. Hemminki E, Gissler M, Toukomaa H. Exposure to female hormone drugs during pregnancy: effect on malformations and cancer. *British journal of cancer* 1999; **80**(7): 1092-7.
172. Kullander S, Kallen B. A prospective study of drugs and pregnancy. 3. Hormones. *Acta obstetricia et gynecologica Scandinavica* 1976; **55**(3): 221-4.
173. Mau G. Progestins during pregnancy and hypospadias. *Teratology* 1981; **24**(3): 285-7.
174. Oates-Whitehead RM, Haas DM, Carrier JA. Progestogen for preventing miscarriage. *The Cochrane database of systematic reviews* 2003; (4): Cd003511.
175. Raman-Wilms L, Tseng AL, Wighardt S, Einarson TR, Koren G. Fetal genital effects of first-trimester sex hormone exposure: a meta-analysis. *Obstetrics and gynecology* 1995; **85**(1): 141-9.
176. Resseguie LJ, Hick JF, Bruen JA, Noller KL, O'Fallon WM, Kurland LT. Congenital malformations among offspring exposed in utero to progestins, Olmsted County, Minnesota, 1936-1974. *Fertility and sterility* 1985; **43**(4): 514-9.
177. Yoon BH, Romero R, Moon JB, et al. Clinical significance of intra-amniotic inflammation in patients with preterm labor and intact membranes. *American journal of obstetrics and gynecology* 2001; **185**(5): 1130-6.
178. Northen AT, Norman GS, Anderson K, et al. Follow-up of children exposed in utero to 17 alpha-hydroxyprogesterone caproate compared with placebo. *Obstetrics and gynecology* 2007; **110**(4): 865-72.
179. de Carvalho MH, Bittar RE, Brizot Mde L, Bicudo C, Zugaib M. Prediction of preterm delivery in the second trimester. *Obstetrics and gynecology* 2005; **105**(3): 532-6.
180. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**(8476): 307-10.
181. Jones M, Dobson A, O'Brian S. A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *International journal of epidemiology* 2011; **40**(5): 1308-13.

182. Kundel HL, Polansky M. Measurement of observer agreement. *Radiology* 2003; **228**(2): 303-8.
183. Greco E, Lange A, Ushakov F, Calvo JR, Nicolaides KH. Prediction of spontaneous preterm delivery from endocervical length at 11 to 13 weeks. *Prenatal diagnosis* 2011; **31**(1): 84-9.
184. Retzke JD, Sonek JD, Lehmann J, Yazdi B, Kagan KO. Comparison of three methods of cervical measurement in the first trimester: single-line, two-line, and tracing. *Prenatal diagnosis* 2013; **33**(3): 262-8.
185. Andrade KC, Bortoletto TG, Almeida CM, et al. Reference Ranges for Ultrasonographic Measurements of the Uterine Cervix in Low-Risk Pregnant Women. *Revista brasileira de ginecologia e obstetricia : revista da Federacao Brasileira das Sociedades de Ginecologia e Obstetricia* 2017; **39**(9): 443-52.
186. Brandao RS, Pires CR, de Souza E, et al. Magnetic resonance imaging vs. transvaginal ultrasound for cervical length assessment in the second half of pregnancy. *Ultrasound in medicine & biology* 2010; **36**(4): 571-5.
187. Burger M, Weber-Rossler T, Willmann M. Measurement of the pregnant cervix by transvaginal sonography: an interobserver study and new standards to improve the interobserver variability. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 1997; **9**(3): 188-93.
188. Furtado MR, Pires CR, Araujo Junior E, De Souza E, Nardozza LM, Moron AF. Transvaginal grey scale histogram of the cervix at 20-25 weeks of pregnancy. *The Australian & New Zealand journal of obstetrics & gynaecology* 2010; **50**(5): 444-9.
189. Goya M, Pratcorona L, Higueras T, Perez-Hoyos S, Carreras E, Cabero L. Sonographic cervical length measurement in pregnant women with a cervical pessary. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2011; **38**(2): 205-9.
190. Pruksanusak N, Sawaddisan R, Kor-Anantakul O, Suntharasaj T, Suwanrath C, Geater A. Comparison of reliability between uterocervical angle and cervical length measurements by various experienced operators using transvaginal ultrasound. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2018; **Sept**: 1-8.
191. Stein W, Hellmeyer L, Schmidt S, Tekesin I. Intraobserver and interobserver reliability of transvaginal cervical length measurements and quantitative ultrasound tissue characterization of the cervix in the second and third trimester of pregnancy. *Ultraschall in der Medizin (Stuttgart, Germany : 1980)* 2011; **32 Suppl 2**: E169-74.
192. Banos N, Perez-Moreno A, Julia C, et al. Quantitative analysis of cervical texture by ultrasound in mid-pregnancy and association with spontaneous preterm birth. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2018; **51**(5): 637-43.
193. Thorlacius H, Toth E. [Implementation of colorectal cancer screening in Sweden]. *Lakartidningen* 2018; **115**.
194. Gates TJ. Screening for cancer: evaluating the evidence. *American family physician* 2001; **63**(3): 513-22.
195. Hellquist BN, Duffy SW, Abdsaleh S, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. *Cancer* 2011; **117**(4): 714-22.

196. Campbell S. Universal cervical-length screening and vaginal progesterone prevents early preterm births, reduces neonatal morbidity and is cost saving: doing nothing is no longer an option. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2011; **38**(1): 1-9.
197. Cahill AG, Odibo AO, Caughey AB, et al. Universal cervical length screening and treatment with vaginal progesterone to prevent preterm birth: a decision and economic analysis. *American journal of obstetrics and gynecology* 2010; **202**(6): 548.e1-8.
198. Einerson BD, Grobman WA, Miller ES. Cost-effectiveness of risk-based screening for cervical length to prevent preterm birth. *American journal of obstetrics and gynecology* 2016; **215**(1): 100.e1-7.
199. Werner EF, Hamel MS, Orzechowski K, Berghella V, Thung SF. Cost-effectiveness of transvaginal ultrasound cervical length screening in singletons without a prior preterm birth: an update. *American journal of obstetrics and gynecology* 2015; **213**(4): 554.e1-6.
200. Werner EF, Han CS, Pettker CM, et al. Universal cervical-length screening to prevent preterm birth: a cost-effectiveness analysis. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2011; **38**(1): 32-7.
201. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bulletin of the World Health Organization* 2008; **86**(4): 317-9.

Tove Jansson. *Muumipapan urotyöt. suom* Laila Järvinen, Schildts 1950

”Myrskyt, ah! Niiden merkitys lienee siinä, että jälkeenpäin näkee auringonnousun. Karttahytti saisi uuden kultanupin. Join kahvini ja olin tyytyväinen.

Ja nyt kääntyi lehti, elämässäni oli alkamassa uusi luku. Maata näkyvissä, suuri yksinäinen saari keskellä merta! Vieraan rannikon ylpeä ääriviiva!”

Tove Jansson. *Muminpappans Bravader skrivna av honom själv*. Schildts 1950

“Stormar ack! Meningen med dem är antagligen att man ska få en soluppgång efteråt. Navigationshyttens skulle få en ny guldknopp. Jag drack mitt kaffe och var nöjd.

Och nu vände sig bladet, jag närmade mig ett nytt kapitel i mitt liv. Land föröver, en stor ensam ö mitt i havet! Den stolta silhuetten av en främmande kust!”