Symphysis Fundus Measurements for Detection of Intrauterine Growth Retardation

EVA BERGMAN
Dissertation presented at Uppsala University to be publicly examined in Rosénsalen, Ingång 95-96, Akademiska sjukhuset, Uppsala, Friday, May 21, 2010 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in Swedish.

Abstract

A case-control study was performed to evaluate the Swedish population-based symphysis fundus (SF) reference curves. The study included 242 small for gestational age (SGA) neonates (169 term and 73 preterm infants) as cases and 296 non-SGA infants as controls. Two Swedish SF curves were evaluated. In term pregnancies they showed a sensitivity of 32 % and 51 % and a specificity of 90 % and 83 %, respectively, at a cut-off level of < -2 SD from the mean according to the SF reference curve. The sensitivity for SGA was higher in preterm pregnancies (49 % and 58 %, respectively) and the first alarm below – 2 SD was noted before 32 weeks in 37 % and 43 % of the preterm pregnancies, respectively. (Study I)

A study of self-administered SF measurements was designed to achieve more regular and frequent SF measurements. Thirty-three women with singleton, ultrasound dated pregnancies performed SF measurements on average 14 weeks from gestational week 20 to 25 until delivery. Self-administered SF measurements were higher and had higher variance than midwives’ measurements. Four consecutive SF measurements on each occasion can compensate for higher variance. Reliable self-administered SF measurements can be obtained. (Study II)

Self-administered SF measurements from 191 women were used to construct absolute and relative SF growth references. The influence of fetal sex, maternal obesity and parity was assessed in regression models. The InSF growth was statistically influenced by maternal obesity, and a borderline significance was recorded for fetal sex and parity. Statistical analysis and graphical displays show no evidence that the relative InSF growth should be dependent on these variables. (Study III)

To improve detection of infants with intrauterine growth restriction (IUGR) rather than SGA a new statistical model (the SR method) was used. The SR method was evaluated with SF measurements from 1122 pregnant women. The sensitivity for neonatal morbidity and SGA was low, between 6 and 36 % for SGA (< -2SD). Neonates classified as SGA (< -2SD and < 10th percentile) had increased morbidity compared with the total study group. Neonates suspected to be SGA before delivery by the population-based SF measurement method had lower morbidity than those not suspected. The SR method was found not to improve detection of fetuses with increased morbidity or SGA neonates in this study. Better screening methods to detect IUGR and SGA prior to delivery are needed. (Study IV)

Keywords: symphysis-fundus measurements, small for gestational age, intrauterine growth retardation, fetal growth, self-administered, relative growth, screening method, statistical surveillance, fetal surveillance

Eva Bergman, Obstetrics and Gynaecology, Uppsala University, SE-75185 Uppsala, Sweden

© Eva Bergman 2010

ISSN 1651-6206
urn:nbn:se:uu:diva-122256 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-122256)
Dedications

When you know a thing, to hold that you know it;
and when you do not know a thing, to allow that you do not know it
- this is knowledge.

Confucius.
Chinese philosopher & reformer
(551 BC - 479 BC)

To my beloved family
This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


*Reprints were made with permission from the respective publishers.*
Contents

Introduction ................................................................................................... 11
General background ................................................................................. 11
Symphysis fundus measurements ............................................................. 11
Normal intrauterine growth and growth rate ............................................ 13
Small for Gestational Age (SGA) ............................................................ 13
Intrauterine Growth Retardation (IUGR) ................................................. 14
Uteroplacental insufficiency ..................................................................... 15
Screening methods ................................................................................... 16
Statistical methods; cross sectional and longitudinal methods ................. 17
Shiryaev Roberts method (SR method) .................................................... 17

AIMS OF THE PRESENT STUDY ............................................................. 19
The specific aims of this thesis were:....................................................... 19

MATERIAL AND METHODS .................................................................... 20
Population and study design ..................................................................... 20
Study I: Symphysis fundus measurement for detection of small for gestational age pregnancies. ................................................................. 20
Study II: Self-administered measurement of symphysis fundus heights ................................................................................................. 21
Study III: Relative growth estimated from self-administered symphysis fundal measurements. ............................................................. 22
Study IV: Symphysis fundus measurements used with a novel statistical method for detection of intrauterine growth retardation; a clinical evaluation ................................................................. 23

Statistical Analyses ....................................................................................... 25
Study I ...................................................................................................... 25
Study II ..................................................................................................... 25
Study III ................................................................................................... 25
Study IV ................................................................................................... 26

RESULTS ..................................................................................................... 27
Study I: Symphysis fundus measurement for detection of small for gestational age pregnancies. ................................................................. 27
Study II: Self-administered measurement of symphysis fundus heights ................................................................................................. 27
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGA</td>
<td>appropriate for gestational age (birth weight within +/- 2 SD)</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the curve</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index (kg/m²)</td>
</tr>
<tr>
<td>BPD</td>
<td>biparietal diameter</td>
</tr>
<tr>
<td>CTG</td>
<td>cardiotocography</td>
</tr>
<tr>
<td>GA</td>
<td>gestational age</td>
</tr>
<tr>
<td>GW, gw</td>
<td>gestational week</td>
</tr>
<tr>
<td>HPL</td>
<td>human placental lactogen</td>
</tr>
<tr>
<td>IUFD</td>
<td>intrauterine fetal death</td>
</tr>
<tr>
<td>IUGR</td>
<td>intrauterine growth retardation</td>
</tr>
<tr>
<td>K curve</td>
<td>Kieler SF reference curve</td>
</tr>
<tr>
<td>lnSF</td>
<td>the natural logarithm of symphysis fundus measurement</td>
</tr>
<tr>
<td>LGA</td>
<td>large for gestational age (birth weight &gt; + 2 SD from mean)</td>
</tr>
<tr>
<td>LMP</td>
<td>last menstrual period</td>
</tr>
<tr>
<td>MBR</td>
<td>The Swedish Medical Birth Register</td>
</tr>
<tr>
<td>ROC curve</td>
<td>Receiver Operating Characteristic curve</td>
</tr>
<tr>
<td>S curve</td>
<td>Steingrimsdottir SF reference curve</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SF</td>
<td>symphysis fundus</td>
</tr>
<tr>
<td>SF method</td>
<td>the method by which SF measurements are compared with population-based references</td>
</tr>
<tr>
<td>SGA</td>
<td>small for gestational age (birth weight &lt; - 2 SD from mean)</td>
</tr>
<tr>
<td>SR method</td>
<td>Shiryaev-Roberts method</td>
</tr>
<tr>
<td></td>
<td>The statistical method derived by Shiryaev and Roberts by which SF measurement data are analysed for detection of deviant growth</td>
</tr>
<tr>
<td>U-estriol</td>
<td>urinary estriol</td>
</tr>
<tr>
<td>W curve</td>
<td>Westin SF reference curve</td>
</tr>
</tbody>
</table>
Introduction

General background
Fetuses and infants small for gestational age (SGA) and in particular those with intrauterine growth retardation (IUGR) have increased risks of morbidity and perinatal mortality (1-3). SGA is defined as a birth weight below a certain limit compared with a population-based reference curve, while intrauterine growth retardation (IUGR) is defined as a failure to reach the genetic growth potential and always implies pathological growth (1, 4). Since fetal growth retardation is difficult to assess, SGA is often used as a proxy for IUGR (1, 2). Antenatal identification of SGA and IUGR fetuses improves perinatal outcome (2, 5-7), and detection of SGA/IUGR prior to delivery is of importance.

In Sweden mainly two methods are used for antenatal detection of SGA fetuses – population-based symphysis fundus (SF) reference curves (here called the SF method) as a screening tool and fetal weight estimates by ultrasound as a screening and diagnostic tool (8-10). However, both methods have a rather low sensitivity for antenatal detection of SGA fetuses. The sensitivity for detecting SGA using SF measurements is reported to be between 16-84% (11-14). The corresponding figure for ultrasound screening is reported to be 50-80% (2, 10, 15, 16). Thus improvements are needed.

Symphysis fundus measurements
The first Swedish population-based reference curve for SF measurements was introduced as a screening test for deviant growth by Westin in 1977 (17). SF measurements were compared with and found superior to maternal weight gain, girth measurements, human placental lactogen (HPL) and urinary estriol (U-estriol) in finding both accelerated and retarded fetal growth. The reference curve was based on pregnancies dated from the last menstrual period (LMP). The SF method by Westin as well as other SF growth charts were tested in the 1980s and proved useful (12-14). All these studies point out that SF measurement is a simple, inexpensive, harmless and widespread screening method.

Instructions on SF measurements according to Westin say that: “The patient should be supine. Legs should be straight otherwise the pubic
symphysis moves upwards. The uterus should be relaxed and the bladder empty. The measurement should be performed along the longitudinal axis of the uterus defined both in the frontal and lateral projections. During the third trimester, measurements should be performed along the longitudinal axis of the fetus whereby fetal crown-rump length will be reflected” (17).

The measured SF height is compared with a cut off level in the population-based reference curve. This method of analysing the result only takes the last SF measurement into account when comparing the measurements to a chosen cut-off level.

The reference curve by Westin was in use until 1996 in Sweden. After the establishment of routine ultrasound for estimation of gestational age, the curve by Westin was replaced by a new one presented by Steingrimsdottir et al constructed from ultrasonically dated pregnancies (8). Moreover, two other Swedish population-based curves for ultrasonically dated pregnancies have been published (18, 19), but have never been applied in clinical practice. The effectiveness of the three ultrasound based reference curves as a screening instrument for detection of SGA fetuses has never been tested previously.

There are some known drawbacks associated with SF measurements. The low sensitivity has been criticised (11, 20, 21) as well as the high interindividual measurement differences (22, 23). For optimal results the SF measurements should be performed by the same person and intervals between measurements ought to be regular and not too protracted. In addition the SF method has been criticized for aiming at detection of SGA fetuses. The primary goal should be to detect IUGR fetuses, which are those at highest risk of adverse outcomes (24-26).

A key issue in constructing effective surveillance methods is the availability of data. The more data available (SF measurement occasions), the better the ability to detect the condition of interest (deviant growth). Secondly, the data has to be analysed according the best available method. Thus, improvements of the SF method should involve an early start as well as closer and more regular measurements and might be improved by using a new statistical method to analyse the SF data (27).

In this thesis self-administered SF measurements are introduced to give access to earlier and more frequent evaluation of fetal growth, which might improve detection of fetuses with deviant growth. Self-administered SF measurements give the pregnant women an added responsibility that could cause some extra anxiety. However, self monitoring is already performed in high risk pregnancies and self-administered tests are commonly used and well accepted among patients with chronic diseases such as diabetes, heart failure and asthma (28, 29).
Normal intrauterine growth and growth rate

Fetal growth could be divided into three phases. During the first 16 weeks of pregnancy growth is due to rapid cellular proliferation. The first phase is mainly genetically influenced and fetal weight and growth will most strongly correlate to gestational age. From 16-32 weeks of gestation fetal growth is dependent on both cellular proliferation and cell hypertrophy. The last trimester is a phase of cellular hypertrophy and is characterised by a rapid increase in cell size. The second and third phases are influenced by both genetic and environmental factors and fetal weight will have a wider distribution (30, 31).

Since fetal growth is fairly equal between fetuses in early pregnancy, ultrasound examination before 20 weeks of gestation is considered the best way to determine pregnancy length and is now the gold standard for dating pregnancies (32, 33). Correct dating of pregnancies is a prerequisite for the evaluation of fetal weight by gestational age and fetal growth later in pregnancy which makes it a base for obstetric management (4, 9, 34, 35).

Previously, neonatal birth weight curves were based on mean birth weight for gestational age-matched neonates. These birth weight curves did not represent the intrauterine population, especially not in the preterm period where growth retardation in neonates is more common (36, 37). Since 1996 Swedish population-based birth weight curves are based on fetal weight estimation by ultrasound, corrected by fetal gender (38). This new weight reference will classify more neonates born premature as SGA. Birth weight curves based on fetal weight estimation by ultrasound is accepted as a more representative way to estimate fetal weight in the normal population than birth weight in premature neonates (30, 35). Aside from gestational age, fetal gender is the most important individual factor influencing fetal weight. Fetal weight is influenced by gender as early as 15 weeks of gestation (39).

Small for Gestational Age (SGA)

It has long been known that SGA fetuses and infants have an increased perinatal morbidity and mortality (1).

However, the definition of SGA varies in different countries and in different studies. SGA is defined as a birth weight below a certain limit compared with a population-based reference curve. SGA only defines the proportion of the population with the lowest birth weight.

In Sweden the definition of SGA is a birth weight below -2 SD from the mean for gestational age and gender (38), and corresponds to the 2.5th percentile. In English literature the 10th or 5th percentile is most often used as a cut-off limit to define SGA (which corresponds to the lowest 10 or 5 % of the population). SGA comprises both genetically small fetuses and fetuses
with pathologically impaired growth. However, as many as 90 % of the SGA infants demonstrate catch up growth during their first 2 years of life (40), indicating a high proportion of IUGR in the SGA group.

Previously about 3.4 % of all newborns in Sweden were classified as SGA according to the Swedish Medical Birth Register (41). During the last 10 years the incidence has decreased to 2.5 %. The new fetal weight reference curve introduced in 1996, based on intrauterine fetal weight estimated by ultrasound, might have influenced the SGA incidence (38). Moreover, the increase in maternal BMI might contribute to the decrease in the SGA incidence. However mean birth weight has not increased during this period.

Intrauterine Growth Retardation (IUGR)

IUGR is defined as a failure to reach the genetic growth potential and always implies pathological growth (4). There is still not a unanimous definition of IUGR (42).

The risk of perinatal morbidity and mortality is more pronounced for IUGR than for SGA infants (1, 3, 43, 44).

To recognise IUGR the individual growth potential or the individual growth velocity has to be evaluated. Since IUGR fetuses are at higher risk of adverse pregnancy outcome (25, 26) there is a need to develop methods for detection of IUGR.

Several attempts have been made to identify IUGR fetuses through both SF measurements and ultrasound weight estimation (45-47) as well as through Doppler ultrasound (7).

One way to approach this problem is to construct individual reference curves. Such “customised” SF and birth weight curves, where maternal weight, length, parity, ethnicity and fetal gender are taken into account, have been constructed and used in the UK as well as in several other countries. Customised reference curves have improved detection of fetuses with adverse pregnancy outcomes in need of neonatal care and are presumably better at identifying SGA neonates that are growth retarded (1, 25, 26, 48).

Neonates classified as SGA by customised curves are believed to more often be IUGR than neonates classified as SGA by population-based curves. A majority of those classified as SGA by population-based reference curves will also be classified as SGA by customised reference curves. About 10-30 % of the SGA population will be reclassified as non SGA by customised curves (25, 26, 44). A model of a classification system is presented in Fig 1.

Notably, however, is that neonates reclassified as IUGR by customised curves but not SGA by population-based curves have the highest risk of stillbirth and adverse pregnancy outcome (44). According to a previous study this increased risk was mainly correlated to reclassification of preterm
infants who have a high morbidity and mortality (37). The same change has been demonstrated by converting to birth weight references based on ultrasound estimated fetal weight, instead of birth weight references based on actual neonatal birth weight (49).

![Figure 1. SGA classified by A population based, B both and C customised birth weight standards.](image)

One drawback with the customised method is that fetal gender, which is often unknown, is the single most important factor for the estimation of fetal growth (39, 45, 50, 51).

Attempts have been made to use growth velocity to identify IUGR. Estimation of growth velocity always necessitates repeated evaluations of fetal weight or other biometric parameters (52, 53). Conditional centile references for weight estimation by ultrasound have also been developed (47).

However, all these methods have limited effectiveness for antenatal detection of IUGR in low risk populations (2, 54).

**Uteroplacental insufficiency**

Uteroplacental insufficiency is thought to be the major cause of IUGR (80-90%) (55, 56). A model of fetal deterioration for this condition has been derived (57), from which clinical surveillance is designed.

When exposed to uteroplacental insufficiency the fetus shows compensatory mechanisms. Firstly, the growth decreases, which can be seen as reduced fetal weight. The impaired efficiency of the placenta can be reflected as reduced blood flow in the umbilical artery and the uterine artery. Gradually the blood flow within the fetus will be redistributed. Decreased
amniotic fluid volume reflects reduced blood flow to the kidneys and decreased urinary production. The blood flow to the brain, examined in the medial cerebral artery (MCA) will primarily increase, but will return to normal if the fetus is severely compromised. One of the final effects in the fetus can be reflected as a changed blood flow pattern in the ductus venosus. Finally, pathological CTG findings appear before intrauterine fetal death occurs (57, 58).

Screening methods

A screening method is not the same as a diagnostic test. A screening test is used in a low risk population with no symptoms of disease. Selective screening can be used for a preselected high risk population.

Since screening tests are used in healthy populations without symptoms of disease there are higher demands on safety and efficacy. The chance to do good must exceed the risk of harm (59, 60). The efficacy of the test can be measured from the ability to detect the disease (sensitivity) and the ability to identify healthy persons (specificity). For every test there is usually a trade off between sensitivity and specificity. A change in the cut-off limit to reach a higher sensitivity will result in a lower specificity, with a larger amount of false positive results. If the disease is rare, the number of false positive cases will exceed the amount of true positive ones. Most screening tests are used to select a high risk population that will then be offered diagnostic tests or examinations. All examinations and tests will create some anxiety in the examined population (59).

SF measurement is a harmless test with no negative side effects except risks of increased anxiety in case of deviating measurements. The effectiveness of the test is, however, rather low. Since SF measurement is harmless it can be accepted in the population despite a low sensitivity (61-63). Fetal weight estimation by routine ultrasound of the fetus in the third trimester has been suggested as a way to detect SGA fetuses (64). However, in a Swedish study the sensitivity for detection of SGA fetuses with such a program was only 54 % (2). The ultrasound method is expensive and requires technologically advanced systems as well as skilled personnel (2). Repeated measurements are required to detect IUGR (4) potential biological side effects from ultrasound examinations have to be considered when proposing such examinations as a screening method in a healthy population (59-63). In addition, from a global perspective repeated ultrasound examinations are not available for most pregnant women. Accordingly, after refinement of the method, symphysis fundus measurements might be applicable as a screening test and weight estimation by ultrasound as a diagnostic method in cases of suspected deviant growth (9, 10, 31, 65).
Statistical methods; cross sectional and longitudinal methods

Statistical surveillance or statistical process control (SPO) was first used in industrial settings to control the quality of the manufactured products (27).

Statistical methods are the means by which we analyse and interpret data from the process under study. Different data can and should be statistically analysed by different methods depending on the circumstances of the study.

Cross sectional data are collected at the same time from different individuals in a specific population. Cross sectional data are suitable for construction of references where the measurement data are to be compared with the population (66). For example; cross sectional data can provide a reference for fetal weight.

Longitudinal data are collected repeatedly from the same individuals over time (67). Longitudinal data collection and analyses are useful when handling processes changing over time, for example fetal growth (68).

Each scientific study and clinical evaluation needs to consider which statistical methods best serve the purpose of the situation. In this thesis the SR method, a longitudinal statistical method for detection of IUGR (i.e. impaired fetal growth) is introduced and evaluated (27).

Shiryaev Roberts method (SR method)

To improve the SF method we have previously developed a statistical surveillance method for deviating SF measurements based on the classical statistical method by Shiryaev-Roberts (here called the SR method). The SR method is a statistical method based on publications from Shiryaev (69) and Roberts (70). The SR method is a (conditional) likelihood ratio method.

More efficient handling of the information in the frequent SF measurements could lead to an improved method for detection of IUGR. In the SR method the deficient growth velocity, expressed as relative growth, is used to detect IUGR and this performs better in a theoretical model than population-based SF measurements as they are used in Sweden today (27). The statistical theory stipulates that the use of the SR method could lead to earlier detection time and fewer false alarms compared with the traditional SF method used today.

In the setting of this thesis the SR method is modulated for optimal detection of IUGR from weekly self-administered SF measurements (27). In the SR method all SF measurements performed during a pregnancy are taken into account, although recent SF measurements are given more weight than older ones.

The SF measurements are first transformed to the natural logarithm (lnSF) for distributional purposes. In the next step relative growth is
calculated by subtracting lnSF from one week with lnSF from the previous week. This step diminishes the influence of individual variables such as maternal BMI, parity and fetal sex and give each pregnancy its own individual reference curve. Subsequently, relative growth is compared with the reference of expected relative growth (derived in study III) when analysed using the *SR method*.

In the studies of this thesis the alarm statistics in the *SR method* is optimised to give best chance of an alarm when a total stop of SF growth occurred.
AIMS OF THE PRESENT STUDY

The overall purpose of this thesis was to evaluate and develop SF measurements as a screening instrument for detection of SGA or preferably IUGR.

The specific aims of this thesis were:

- To evaluate the three Swedish population-based SF reference curves constructed from ultrasound dated pregnancies (Study I).

- To evaluate if pregnant women can perform SF measurements by themselves to obtain SF measurement data more often and regularly and to compare self administered SF measurements with measurements performed by midwives (Study II).

- To construct a relative growth reference curve from self-administered SF measurements for subsequent use in the SR method (Study III).

- To evaluate the ability of the SR method to detect IUGR assessed as neonatal morbidity (Study IV).
MATERIAL AND METHODS

Population and study design

Study I:
Symphysis fundus measurement for detection of small for gestational age pregnancies.

This study was designed as a case-control study. All singleton SGA infants born between 1993 and 1997 in the county of Uppsala, Sweden were included as cases after identification from the Swedish Medical Birth Register (MBR) (71). SGA was defined as a birth weight below 2 SDs from the mean according to gender-specific Swedish birth weight charts (38). The controls were term (gestational age ≥ 37 completed weeks), singleton, non-SGA infants born in Uppsala during the same period. Mothers from outside the county are expected to represent a high rate of pregnancy complications, and were excluded from the study.

The Swedish Medical Birth Register MBR was used to identify the SGA and control infants and to obtain mortality data. All other data on maternal and neonatal characteristics as well as SF measurements were retrieved from the maternal and neonatal medical records.

The MBR identified 308 SGA infants and 300 controls. Among these, 242 SGA cases (of which 169 (70 %) were term and 73 (30 %) preterm pregnancies), and 296 controls were included in the study.

For comparison of neonatal birth weight we used birth weight ratios since gestational age is the most important factor influencing neonatal weight. The birth weight ratio was calculated as: birth weight/expected birth weight by gestational age.

Analyses for sensitivity and specificity of the three Swedish SF curves concerning detection of SGA pregnancies were studied. The reference curve constructed by Steingrimsdottir et al (S curve) (8) and Kieler et al (K curve) (19) include data on mean SF values as well as standard deviations. SDs were not available for the curve presented by Håkansson et al. (18). Therefore, this reference curve was excluded from the analyses. For the S and K curve 2 SDs correspond to approximately 2.8 cm in the 25th week of gestation and 3.2 cm in the 40th week. The curve by Håkansson is situated in between the curves by Steingrimsdottir and Kieler.
Sensitivity in detecting SGA was tested in the SGA pregnancies for cut-off limits from 1 to 4 SDs below the means with intervals of 0.25 SDs, and with no increase in three consecutive SF measurements. Specificity was presented from the controls of which all were term pregnancies. Receiver operating characteristic (ROC) curves were constructed for term pregnancies.

In the preterm SGA group gestational age at the first SF alarm signal for different cut-off levels was recorded.

**Maternal and neonatal characteristics**

The mothers of the term SGA infants were shorter, more often primiparous, heavier smokers, and more often diagnosed with hypertension compared to the control mothers. Term SGA infants had higher mortality, were more often in need of neonatal care, and had more malformations diagnosed compared with controls.

The mothers of preterm SGA infants were older, taller, and had a lower BMI, fewer were heavy smokers and their hypertension rate was higher compared to mothers of term SGA infants. Preterm SGA infants had the lowest birth weight ratio, the highest mortality and morbidity rates and were more often delivered by operative delivery due to fetal distress.

**Study II:**

Self-administered measurement of symphysis fundus heights.

This study was designed as a prospective longitudinal study. Forty pregnant women from two primary antenatal clinics in Uppsala City were consecutively recruited into the study, during 2003-2004. Six women did not perform SF measurements and one woman lost her tape measures. Data from 33 women were available in the analyses. Inclusion criteria were singleton, ultrasound dated pregnancies. Exclusion criteria were lack of Swedish language skills and anticipated move from the county during pregnancy.

Written and oral instructions on how to perform SF measurements according to Westin (17) were given to all participants by their midwife at the antenatal clinics. The pregnant women used blank paper tape measures and were asked to perform four consecutive SF measurements once a week from gestational week 20-25 until delivery. All women attended an ordinary routine antenatal care program which includes 7 to 9 check ups from pregnancy week 20 until delivery. During the routine antenatal visits SF measurements were performed by the midwife. The midwives utilise the plastic covered fabric tape measures used routinely at the antenatal clinics for their regular SF measurements (open). To evaluate if the blank paper tape measures used by the pregnant women were comparable to the plastic covered fabric tape measures, and to evaluate if SF measurements were
influenced by blinded or open measurements, the midwives additionally performed one SF measurement with the blank paper tape measure (blinded) before their regular measurement (open).

Of the 33 pregnant women included, two discontinued participation after one and three weeks, respectively. Thirty-one women performed SF measurements until gw 34 + 6 or longer for a period of 14 weeks (median, range 1 – 21). All but three performed four SF measurements on every occasion.

The fifteen midwives participating in the study performed on average 6 (median) (range 4 – 8) SF measurements, on each woman during the pregnancies.

Maternal and neonatal characteristics
The women were healthy, had a mean age of 31 years at delivery, a mean height of 166 cm and a mean BMI of 24 at booking. There were no pregnancy complications among the participants and all delivered at term.

All neonates were healthy although two were classified as SGA (38) and four as large for gestational age (LGA). The mean birth weight was 3490 grams (range 2654 to 4600).

Study III:
Relative growth estimated from self-administered symphysis fundal measurements.

The same longitudinal study design as in Study II was applied. Three hundred pregnant women with singleton ultrasound dated pregnancies, from five primary antenatal clinics in Uppsala and Märsta (Stockholm) were consecutively recruited into the study during 2006 - 2007.

Instructions on how to perform self-administered SF measurements from gw 20 until delivery were given according to Westin (17, 72) in the same way as in Study II. For practical reasons the women used specially designed tape measures similar to the one used by midwives. Self-administered SF measurements were recorded in a specifically designed diary and collected after the delivery.

Since the antenatal visit in gw 20 was excluded from the Swedish antenatal programme during the study period, SF measurement data obtained before gw 25 were excluded from the final analyses.

Eighty seven women did not perform any SF measurements or failed to deliver their measurement diaries. One woman moved abroad and was lost to follow up. The 212 pregnant women (71 %) performed self-administered SF measurements on average 14 weeks (range 1-23) during pregnancy.

For the construction of a relative growth reference curve with the purpose of finding IUGR, 19 pregnancies with factors known to negatively influence
fetal growth (hypertension, smoking and preterm delivery) were excluded. Additionally, two women who did not perform measurements after gw 25 were excluded. SF measurements from 191 women were included for the construction of the absolute and the relative growth reference curve, and for comparison with midwives’ measurements.

Maternal, neonatal and delivery data were obtained from the medical records. SGA was defined as a neonate weighing < − 2 SD and large for gestational age (LGA) as > +2 SD from the mean according to Swedish birth weight standards for gestational age and gender (38).

Maternal and neonatal characteristics
The women had a mean age of 31 years at delivery, a mean height of 168 cm and a mean BMI of 24 (kg/m²) at booking. Ninety one percent of the women delivered at term.

Except for lower asthma incidence among non performers (2.3 % vs 10.4%, p=0.02), no statistically significant differences were found for maternal characteristics and neonatal outcomes between performers and non performers of self-administered SF measurements.

The neonates had a mean birth weight of 3550 grams (range 648 to 5200). Four neonates were classified as SGA, of which two were born preterm. Six neonates were classified as large for gestational age (LGA) (38). One intrauterine fetal death (with normal fetal weight) in pregnancy week 32 + 6 was recorded.

Study IV:
Symphysis fundus measurements used with a novel statistical method for detection of intrauterine growth retardation; a clinical evaluation.

The same longitudinal study design as in study III was applied in a larger setting. 1888 women from 27 Swedish antenatal clinics in 4 counties (Uppsala, Stockholm, Västerås and Gävle), were consecutively recruited from 2006 to 2009. The same instructions on how to perform self-administered SF measurements were given as in study III.

Two thousand one hundred and fifty seven pregnant women were asked to participate in the study, and 1888 accepted enrolment. Fourteen women were excluded due to missing personal identification numbers, preventing access to their medical records. Thus, data from 1874 women were registered. Six hundred and thirty five women (33.9 %) did not deliver SF measurement data and 117 (6.3 %) performed SF measurements less than three pregnancy weeks after gw 24. Data were available on 1122 (61.0 %) women for the final analysis according to the SR method.
Studied predictors were weekly self-administered SF measurements and SF measurements performed by midwives according to the Swedish antenatal care program. Neonatal morbidity was assessed as a primary outcome as a proxy for IUGR. Neonatal morbidity was estimated as Apgar score <7 at five minutes, low arterial pH in the umbilical vessels, respiratory distress and admission to neonatal care (length of stay expressed as mean days/infant and number of infants in neonatal care). In addition we assessed the constructed category “any neonatal morbidity”, which included any of the above given morbidity measures or severe congenital malformations, neonatal jaundice and severe infection.

Secondary outcome was birth weight defined as SGA, with the limits set at < -2 SD and < 10th percentile according to Swedish birth weight standards for gestational age and gender (38).

**Maternal and neonatal characteristics**

Maternal characteristics from the 1122 women show a mean age of 31 years at delivery and a mean body mass index (BMI) of 24 kg/m² at booking. Delivery at term was recorded for 91 % of the women. 2.1 % were smokers compared to 5.6 % in the group not performing self-administered SF measurements (p < 0.001). For all other maternal characteristics no statistically significant differences were found between those who performed and those who did not perform SF measurements.

Among the 1122 women 53 (4.6 %) delivered preterm and 82 (7.2%) had an operative delivery due to fetal distress (ODFD).

Neonatal data show 1122 live born infants with a mean birth weight of 3560 grams (range 980 to 5310). No intrauterine or neonatal deaths occurred in the study group (compared to 4 intrauterine deaths among the 635 women who did not perform SF measurements).

Nineteen cases (19/1122, 1.7 %) were classified as SGA (< - 2 SD) of which 2 (10.5 %) were born preterm. Correspondingly, 91 (8.4 %) cases were classified as SGA according to the definition < 10th percentile of which 10 (11 %) were born preterm.

Of the infants 30 (2.6 %) had respiratory distress, 11 (1.0 %) hypoglycaemia, 14 (1.2 %) an Apgar score < 7 at 5 minutes, 4 (0.4 %) a pH < 7.01 in the umbilical artery and 69 (3.7 %) were admitted to neonatal care.

Studies I – IV were approved by the regional ethics committee in Uppsala
Statistical Analyses

Study I
The reference curves and the SDs for each week were obtained from the authors (8, 19). The SF values for each mother were compared with cut-off levels from -1 to -4 SDs for two population-based SF measurement reference curves, using the Statistical software program package SAS 8.2 (SAS Institute, Cary, North Carolina, USA). Sensitivity and specificity were calculated for the different cut-off levels.

Study II
The results from the SF measurements were analysed by using the Statistical software program package SAS 9.1 (SAS Institute, Cary, North Carolina, USA).

For the analyses of intra- and interindividual variation a mixed two-level model with an individual specific random intercept was applied to the logarithm of the SF data (67). The individual intercepts describe the interindividual variation, and the corresponding residuals describe the intraindividual variation. All the regression analyses as well as the comparison of the variances are based on logarithmic values, while other results are presented in the original scale. The confidence intervals include 95 % of values and stated significances are at the 5 % level.

Study III
The statistical software program package SAS 9.2 (SAS Institute, Cary, NC, USA) was used for the analyses.

The relation between the natural logarithm of the SF measurements, lnSF, and the gestational age in days was modelled with a linear mixed model for longitudinal data. The natural logarithm was applied to simplify the shape of the curve and the distribution of data for modelling purposes.

The model includes clustering of the four SF measurements per gestational week nested within the pregnancy. Thus the dependencies between the repeated measurements per week and over time within one
pregnancy are built into the statistical model. The model was established by a systematic testing of different fractional polynomial models evaluated in terms of goodness-of-fit (73). In the final model the influence of fetal sex, maternal BMI and parity was assessed.

The relative growth was defined as the lnSF measurement at a chosen week subtracted by the measurement the week before. Due to a limited sample size the influence of covariates was assessed both by statistical testing and graphical methods.

A p-value < 0.05 was considered significant.

Study IV

The statistical software package SPSS Statistics 17.0 was used to analyse maternal and neonatal characteristics, as well as neonatal morbidity data. Independent Student’s T-test was used for comparisons of mean values. The Chi square test, and Fisher’s exact test when appropriate, was used for categorical data. A p-value < 0.05 was considered statistically significant.

SF measurements (self-administered and midwives’) were analysed using the statistical software program package SAS 9.2 (SAS Institute, Cary, NC, USA). The relative growth was defined as the lnSF measurement at a chosen week subtracted by the measurement of the previous week. The relative growth calculated in percent from pervious week was subsequently analysed according to the statistical method of Shiryaev Roberts (27, 69, 70).

The alarm statistics in the SR method was optimised to give best chance of an alarm when a total stop in SF growth occurred.
RESULTS

Study I:
Symphysis fundus measurement for detection of small for gestational age pregnancies.

The sensitivity for term SGA in pregnancies of the S curve was 32 % with -2 SDs as the cut-off value. When we included pregnancies with no increase in three consecutive SF values the sensitivity increased to 60 %. The specificity then decreased from 90 % to 72 %. For the K curve the sensitivity was 51 % and the specificity 83 % for term SGA pregnancies with –2 SDs as the cut-off value. When pregnancies with no increase in three consecutive measurements were added, the sensitivity increased to 65 %, and the specificity decreased to 66 %. ROC curves show the same area under the curve for both the S and K curves in term pregnancies.

For preterm SGA pregnancies the S curve had a sensitivity of 49 % for –2 SDs, which rose to 51 % when the group with no increase was added. For the K curve the corresponding values were 58 % and 59 %, respectively. The first SF measurement below 2 SDs in the preterm SGA group was found before 32 gestational weeks for 37 % of the cases with the S curve and in 43 % with the K curve.

Study II:
Self-administered measurement of symphysis fundus heights.

Individual SF curves constructed from self-administered SF measurements have the same shape as previously constructed population-based Swedish SF reference curves. Variance analyses show that self-administered measurements have an interindividual variance of 0.014 and an intraindividual variance of 0.0054 (logarithmic values). The intraindividual variance for self-administered SF measurements was about 5 times higher compared to the midwives’ open measurements. The variance was found to be constant during pregnancy and did not vary with fetal size.

The midwives’ blinded SF measurements were on average higher than their corresponding open SF measurements. The agreement between blinded and open SF measurements performed by midwives was stable over time.

The difference in individual cases between self-administered measurements and midwives’ open measurements most often pointed in the
same direction. The differences between self-administered measurements performed at gestational week 25, 29, 33 and 37 and midwives’ measurements (measured within +/- 4 days) varied from -6.7 to 10.4 cm (72). Twenty-three of 54 such measurements differed 2 cm or less. The average differences and variations between mean values of self-administered SF measurements and midwives’ open measurements at each gestational week were small and constant during pregnancy. A tendency was found for self-administered SF values to be higher than the corresponding measurements performed by midwives.

**Study III:**  
Relative growth estimated from self-administered symphysis fundal measurements.

**SF measurement data**

The self-administered SF measurements were on average situated on a higher level and had a higher variance than SF measurements performed by midwives. The lnSF of the self-administered SF measurements, were found to be fitted by a linear mixed model of the same structure as previously suggested (27) in the model of the SR method and as found in Study II.

The lnSF curves were found to be statistically influenced in terms of level and slope by maternal obesity (defined as BMI ≥ 30, p values 0.003 and 0.002), and border line significant in terms of slope by fetal sex (p-values 0.053 and 0.075) and parity (p-value 0.055 and 0.063).

**Calculating relative growth**

The systematic testing of models to fit the lnSF data gave the following result for the i:th individual at gestational age $t$ in days:

$$\ln SF_i(t) = a_i + b * t + c * \sqrt{t} + e_i(t)$$

where $a_i$ is the individual specific intercept, $b$ and $c$ slope coefficients and $e_i(t)$ is the residual.

The relative growth, $\Delta_i(t)$, was then calculated as:

$$\Delta_i(t) = \ln SF_i(t) - \ln SF_i(t - 1) = b + c * (\sqrt{t} - \sqrt{t - 1}) + \tilde{e}_i(t).$$

Inserting the estimates of the coefficients we obtained the relative growth reference curve as

$$\Delta(t) = -0.0181 * t + 0.666 * (\sqrt{t} - \sqrt{t - 1})$$

A graphical display of the cross-sectional mean SF values and the corresponding relative growth values for lnSF per gestational week is shown (Paper III). Corresponding graphs are shown stratified for fetal sex, maternal
obesity (BMI ≥ 30) and parity. The graphical display shows no evidence that relative growth should be dependent on fetal sex, maternal obesity and parity. Running regression models for relative growth including separate intercept and slope for fetal sex, obesity and parity, respectively, gives p-values > 0.44 supporting the patterns shown in the graphs.

Study IV:
Symphysis fundus measurements used with a novel statistical method for detection of intrauterine growth retardation; a clinical evaluation.

**SF measurement data**
The women performed self-administered SF measurements on average 12.7 weeks after gw 24 with a range from 3 to 18 weeks.

**Predictive ability of the SR method**
ROC curve analysis was performed revealing an area under the curve (AUC) of 50 % for all tested neonatal outcomes, except hypoglycaemia where the AUC was slightly higher.

At the alarm level of -2 SD at any time during the pregnancy, the *SF method* yielded 10.5 % (n=118) test positive women. For comparisons between the *SR* and the *SF method*, the same proportion of the study population (10.5 %, n=118) was chosen for the *SR method*.

Under these circumstances the *SR method* detected 9.8 % of preterm deliveries, 9.9 % of ODFD, 13.8 % of infants with respiratory distress, 36.4 % of infants with hypoglycaemia, 7.1 % of infants with Apgar score <7 at 5 minutes, 6.3 % of infants with pH < 7.01 in the umbilical artery, 6.3 % of any neonatal morbidity and 8.8 % of neonates admitted to neonatal care.

The *SR method* detected 7 cases (36.8 %) of SGA (< -2 SD) and 19 (20.9 %) of SGA (< 10th percentile).

The first alarm from the *SR method* occurred in gw 33.1 (mean).

**Predictive ability of the SF method**
As mentioned above, the *SF method* identified 118 (10.5 %) pregnancies as test positive. The sensitivity was 13.7 % for preterm delivery and 6.2 % for ODFD.

Sensitivity for neonatal morbidity was 13.8 % for respiratory distress, 9.1 % for hypoglycaemia, 7.1 % for Apgar score <7 at 5 minutes, 8.8 % for any neonatal morbidity and 8.8 % for neonates admitted to neonatal care. No test positive pregnancies detected by the SF method resulted in a pH < 7.01 in the umbilical artery.

The *SF method* had a sensitivity of 52.6 % (10/19) for SGA (< -2 SD) and 28.6 % (26/91) for SGA (< 10th percentile).
The first alarm in the SF method occurred in gw 30.0 (mean, range 29-41).

**Morbidity associated with the diagnosis of SGA**

It was more common among neonates defined as SGA to have any neonatal morbidity compared with all the 1122 infants (statistically significant). The outcome variable “any neonatal morbidity” was found in 8 of 19 neonates (42 %) defined as SGA (< -2 SD) compared with 19.9 % in non SGA neonates. Corresponding numbers were 30 (33.0 %) out of 91 SGA (< 10th percentile) compared with 19.2 % for non SGA.

Neonates, in whom SGA was suspected with the *SF method* prior to delivery had, for almost all outcome measures, a tendency toward lower morbidity than test negative neonates.
DISCUSSION

Study I:
Symphysis fundus measurement for detection of small for gestational age pregnancies.

We found that the currently used Swedish SF reference, the S curve (8), had a low sensitivity for detecting SGA. Using a cut-off limit of 2 SDs below the mean, the sensitivity was 32 % and the specificity 90 % for term SGA pregnancies. With lower cut-off values the sensitivity will increase and the specificity decrease. When the cut off level was set to below -1 SD we found a sensitivity of 72 % and a specificity of 71 %. Reducing the specificity from 90 % to 71 % means that another 19 % of the pregnant population will have an extra ultrasound, given that ultrasound is performed to evaluate fetal growth only when growth retardation is suspected.

The K curve (19) is situated about 8 mm higher, and had a higher sensitivity but lower specificity for the corresponding cut-off values.

The sensitivity and specificity patterns were the same for both curves which is shown by ROC curves. No obvious preferential or appropriate cut-off level could be distinguished. Changing the level of the curve or using other cut-off values will not improve the method of population-based SF reference curves.

The sensitivity was higher for the preterm SGA group, 49 % and 51 % for the S and K curve, respectively, indicating that the most severe cases of growth retardation were more often discovered. Our results show that 37 % (S curve) and 43 % (K curve) of the preterm SGA pregnancies were detected before 32 weeks of gestation when – 2 SDs was used as the cut-off value. Since a routine third-trimester ultrasound for detection of SGA pregnancies often is performed at 32 weeks of gestation SF measurements might offer an advantage for detecting the most severe SGA pregnancies compared to a time-fixed ultrasound examination (74).

Major strengths of this study are the external validity and the low risks of selection bias. Our study population was a mixture of urban and rural women and should be representative of large areas of Sweden. The number of cases and controls lost to follow-up was small.

Methods for detecting SGA by means of customised or individualised SF growth curves have been proposed and tested. Such curves take into account
maternal height, weight, parity and ethnicity, and fetal gender (45). A major draw back is that fetal gender, which is the single most important factor influencing SF growth, often is unknown. However, customised reference curves are proven better to detect fetuses with adverse pregnancy outcome and in need of neonatal care (24, 25). Inclusion of maternal weight might reduce detection of SGA pregnancies in underweight women and increase suspicion of SGA in overweight women (37). Ethnicity could only be used in segregated populations. Hence, other methods independent of fetal gender and maternal weight and ethnicity might well be more suitable for detection of SGA and superior for identifying IUGR.

Evaluation of fetal growth velocity could be another way to detect IUGR instead of SGA. Only a few papers have examined and discussed this method (27, 46, 52, 53).

A new methodology for detection of IUGR using repeated SF measurements in combination with a traditional method for statistical surveillance (SR method) (27) might improve detection of these fetuses.

Study II:
Self-administered measurement of symphysis fundus heights.

We found a high level of agreement between self-administered SF measurements and the routine SF measurements performed by midwives. There was a high degree of compliance for four consecutive as well as weekly measurements among the pregnant women and the agreement between midwives’ open and blinded SF measurements was good.

Self-administered SF measurements eliminate interindividual measurement differences and give access to more frequent measurements, which is a limitation of the population-based SF method (22, 23).

In the present study the intraindividual variation for self-administered SF measurements was not influenced by pregnancy length or fetal size. This indicates that pregnant women perform SF measurements in the same way throughout pregnancy and that measurements in early and in late pregnancy have the same reliability.

The given intraindividual variance is based on a regression analysis including all single measurements. A higher variance for self-administered (0.005) compared to midwives’ measurements (0.001) was found. Since the routine for midwives is only one measurement per occasion we chose to present and compare variances estimated similarly for women and midwives knowing that this makes the self-administered measurements look less favorable. However, collecting several self-administered measurements on
each occasion enables the use of mean values for which the variance decreases by the number of measurements.

High compliance among the pregnant women is a major strength of the present study, where almost all self-administered SF measurements were performed according to the study protocol, and 31 of 40 pregnant women completed the study at least until gestational week 34+6.

The population in the catchment areas is similar to that of the general population, which increases the generalisability of the results. The data collected are sufficiently large to make preliminary conclusions about the usefulness of self-administered measurements. However, the size of the study is too small to evaluate the ability of the method to detect SGA fetuses and larger studies in clinical settings are needed to assess this aspect.

For practical reasons and to enable blind measurements the pregnant women used paper tape measures. The difference in the material of the tape measures used by the pregnant women and the midwives complicates the comparisons. The midwives’ measurements with paper tape measures were significantly higher compared with their fabric tape measurements. However, the correlation between their paper and fabric tape measurements was high.

To our knowledge there is only one previously published report on self-administered SF measurements (75). In that study only 60 out of 100 women completed the study compared with 78 % in the present study, and SF measurements were performed as a single measurement once every fortnight. In contrast to our study no comparisons with SF measurements performed by midwives and no analyses of the variance of the measurements were done.

For screening tests, ethical aspects are of great importance since the test mostly involves healthy individuals. All screening tests render some anxiety (60). Self-administered SF measurements give the pregnant women an added responsibility that could cause some extra anxiety. In this study the pregnant women attended the routine antenatal program and had access to the ordinary surveillance program while performing self-administered SF measurements. Self monitoring is already performed in high risk pregnancies and self-administered tests are commonly used and well accepted among patients with chronic diseases such as diabetes, heart failure and asthma (28, 29). We believe that self-administered SF measurements would be acceptable to women if it improves surveillance.

In this type of study there is a risk of including healthy people more interested in their health described as “healthy people bias” as well as an increased number of persons with known increased risk for SGA (59).

Self-administered SF measurements give access to earlier and more frequent evaluation of fetal growth, which might improve detection of fetuses with deviant growth. The findings presented here outline a new way
of combining SF measurement data with an acknowledged statistical surveillance method (27).

Study III:
Relative growth estimated from self-administered symphysis fundal measurements.

The absolute and relative growth reference curve based on self-administered SF measurements follows the structure obtained in an earlier study (27). In line with our previous finding (Study II) we found that self-administered SF measurements were situated at a higher level and had higher variance than measurements performed by the midwives (72). However, when assessing relative growth the differences in level between self-administered and midwives’ SF measurements are of no great concern provided that each woman measures consistently during pregnancy.

The repeated self-administered SF measurements on each occasion compensates for the higher variance. In addition, the low variance in the midwives measurements might not necessarily reflect a higher precision in measuring technique as it could be a sign of a normative behaviour, i.e. adjustment towards the mean.

Fetal sex, maternal obesity in early pregnancy and parity do not appear to influence relative fetal growth. The graphs support the hypothesis that relative growth is independent of these covariates.

A major strength of the present study is that the amount of SF data collected is sufficiently large to construct an absolute and a relative growth SF reference curve.

However, the material collected is not large enough to lend itself to equality analyses in subgroups as assessing the impact of covariates on relative growth. We have previously hypothesised (27) that the dependency on covariates, such as fetal sex, maternal BMI and parity, is less when considering relative growth in SF compared to absolute values. The graphs and the p-values support our initial hypothesis.

The compliance among the pregnant women was 71 %, which we find satisfactory taking into consideration the effort by each woman to perform all measurements. Almost all self-administered SF measurements were performed according to the study protocol as in a previous study (Study II).

Maternal and neonatal characteristics were largely in agreement with characteristics of the women that gave birth during the study period in Sweden and reported to the MBR (41).

No statistically significant differences in maternal or neonatal characteristics were found between the women who performed and those
who did not perform self-administered measurements, except for lower asthma incidence in non performers of self-administered measurements.

The mean birth weight was 49 g higher compared with the Swedish population in 2007. The slightly higher mean birth weight among study participants could be a “healthy people bias” previously discussed in Study II (59, 72).

The catchment areas should be representative of the general population in Sweden, which increases the generalisability of the results.

We believe that compliance can increase to above 71% if the method is shown to improve detection of IUGR.

Study IV:
Symphysis fundus measurements used with a novel statistical method for detection of intrauterine growth retardation; a clinical evaluation.

We performed a longitudinal study assessing the ability of the SR method to detect IUGR neonates, assessed as neonatal morbidity. Neonatal morbidity detected by the SR method was compared with the SF method and the performance by either method was low. The SF method was, as expected, better at detecting SGA neonates.

SGA neonates had a higher morbidity compared with the total study population. Neonates suspected to be SGA by the SF method before delivery had lower morbidity compared with neonates not suspected to be SGA.

As there is no obvious way to define IUGR we used various measures of neonatal morbidity as a proxy for IUGR when assessing the ability of the SR method to detect this condition.

Of these measures, only neonatal hypoglycaemia showed a higher proportion of test positive infants than expected from a random sampling. Transient neonatal hypoglycaemia can be considered a less specific measure, which is rather common even among otherwise healthy infants, but it could also be a sign of intrauterine starvation (76). However, the sensitivity for detection of hypoglycaemia was as low as 36%. It could be that the SR method might prove better in high risk pregnancies.

Relative growth in combination with the SR method is a novel way to analyse SF measurement data based on an optimal statistical theory to give alarm signals for altered growth speed (27). Despite the theoretical advantages, the SR method came out less effective for detection of IUGR than expected. We can see four possible explanations for this.

First of all, the SR method might not have been used in an optimal way. The alarm statistics in the SR method were optimised to give the best chance of an alarm when a total stop in SF growth occurred. The course of events
when intrauterine growth is constrained is not known, but presumably a less abrupt change than a total stop is possible. In future studies the alarm statistics can be modulated to give maximum alarm when a less extreme situation takes place, which, however, requires additional statistical methodological work to be done.

Another explanation could be that the SF measurements using a tape measure are too imprecise for detection of deviant fetal growth, which is illustrated by the low sensitivity of the SF method (77).

Thirdly, growth speed and alterations in growth speed might not be of importance as long as the fetal weight or size is within the “biological optimal window”. Fetal adjustment to alterations in growth might be effective within certain weight limits, and altered growth speed might increase morbidity only when fetal weight is outside these limits.

Finally, despite the clear associations between neonatal morbidity and IUGR, it could be questionable to use morbidity as a proxy of IUGR (1, 4, 25).

To our knowledge this is the first study that has evaluated whether relative fetal growth measured by SF measurements could be used as a method to detect IUGR. Other attempts have been made to optimise detection of IUGR, such as customised SF charts (48) and conditional assessment of fetal growth by ultrasound (46, 47). However serial ultrasound examinations have not yet proved useful as screening instrument in low risk pregnancies (54). Fetal growth speed estimated by ultrasound has been shown to predict emergency caesarean section in a low risk population (78), but not to predict shoulder dystocia in a population of pregnant women with impaired glucose tolerance (79).

There is still a lack of knowledge when it comes to intrauterine fetal growth and growth speed. The greatest challenge is to create improved methods for antenatal detection of IUGR and SGA pregnancies.

Except for a slightly lower prevalence of smokers (2.1 % vs. 7.2 %), hypertension (5.9 % vs. 9 %) and SGA incidence (1.7 % vs. 2.5 %), and a higher mean birth weight (61 g) the women and neonates included in the study were similar to the general pregnant population in Sweden (41). As in study II and III, we interpret our finding as a “healthy people bias”(59, 60).

Although the similarities between our study population and the general pregnant population were more frequent than the dissimilarities it is possible that the lower prevalence of smokers and SGA infants might have decreased our ability to detect IUGR as these factors are closely and inversely related to neonatal morbidity.

The participation and compliance rate were lower than expected, which decreased the power of the study. Of the women asked to participate 61 % accepted and complied with the study criteria of at least three measurements after gestational week 24, which was fewer than we predicted from our previous studies (72, 80).
Since IUGR is highly associated with neonatal morbidity and adverse pregnancy outcome, we used neonatal morbidity as a proxy for IUGR (1, 4, 25). The morbidity measures used are in agreement with measures frequently used in other studies to evaluate neonatal outcome (81). Neonatal morbidity is low in Sweden, which decreases the possibility of finding statistically significant differences. As the proportion demonstrating neonatal morbidity found with the SR method was equivalent to the proportion in the study population our findings contradict the hypothesis that the SR method should detect a higher proportion of IUGR defined as neonatal morbidity.

Since the SF method, in contrast to the SR method, was used in clinical practice to identify deviant fetal growth among study participants the comparison between the two methods are not without bias. Test positive pregnancies identified by the SF method presumably triggered some type of action such as more intensive surveillance or delivery. Accordingly we found a more favourable neonatal outcome in SGA fetuses identified prior to delivery, which is in agreement with previous studies (2, 5). The SGA definition and antenatal detection still has a place in clinical practice.

The Swedish definition of SGA (< -2 SD) can be questioned as the neonatal morbidity was also increased for SGA neonates with a birth weight < 10\textsuperscript{th} percentile. Individually constructed reference curves might be a way forward to improve the antenatal detection of fetuses at risk (46-48).

No cases of intrauterine fetal death (IUFD) occurred in our study population. According to the current Swedish incidence, 3 to 4 cases of IUFD per 1000 deliveries should have been expected. High awareness of pathological fetal growth at the antenatal centres, maybe due to participation in the study, might have resulted in swift intervention when suspicion of SGA occurred and thus affected the result. “Healthy people bias” or a random finding are other explanations.

Clinical approach to SGA and IUGR

SGA and IUGR pregnancies are one of the most common problems in obstetrics. Antenatal detection and correct clinical management improve outcome for these neonates (2, 5, 58).

Only about half of the fetuses are detected prior to delivery and improvement of screening methods would be profitable (2, 45, 77).

In Sweden symphysis-fundus measurements or fetal weight estimation by ultrasound are used as screening methods. Other methods such as uterine artery Doppler and maternal risk factors are also available.

Uteroplacental insufficiency is the cause of IUGR in 80 – 90 % of cases (55). A model of fetal deterioration for this condition has been derived (57) from which clinical surveillance is designed.
Once a screening method or high risk evaluation has raised suspicion of a growth retarded fetus, fetal size can be estimated by ultrasound and compared with a cross sectionally derived population-based reference. Fetal growth (change in fetal size over time) needs a longitudinal approach and at least 2 observations to be evaluated (58). Evaluation of fetal growth with longitudinal methods is increasing, rather than examination of fetal size only. These two methods give different information about the fetus, and the change over time is important for management of the pregnancy (58).

Once the fetus at risk of SGA or IUGR is identified surveillance with ultrasound for fetal weight and growth as well as Doppler examination of blood flow is in place.

Doppler examination with ultrasound has the best predictive value to monitor health in IUGR fetuses (58). Placental dysfunction causes alterations in blood flow in the uterus, placenta and fetus. Impaired fetal growth can be reflected in various Doppler examinations. Blood flow can be examined in the umbilical artery and vein as well as in the uterine artery, where decreased blood flow reflects impaired placental function. Further, Doppler examination of the fetal circulation in the MCA and the ductus venosus reflects fetal compromise. The Doppler results are most informative if they are collected longitudinally, thus enabling evaluation of progress of fetal hypoxia (58).

Today, few options are available for intrauterine treatment. Doppler findings, CTG recordings, amniotic fluid evaluations and fetal growth are used for surveillance with the aim of timing the delivery.

Operative deliveries are common in SGA and IUGR pregnancies. If the fetus is suffering from relative hypoxia before the onset of labour, even normal uterine contractions can lead to fetal distress. An active decision on delivery mode and high awareness of the need of intensive surveillance during labour is probably of great importance for outcome in the individual growth retarded fetus.
GENERAL CONCLUSIONS

• The two tested Swedish SF reference curves based on ultrasound dated pregnancies have low sensitivities for term SGA pregnancies. Sensitivity was higher for the preterm group and SF measurements seem to be better at detecting the most severe cases of SGA. ROC curves reveal that sensitivity and specificity patterns were the same for both tested SF reference curves without any optimal cut-off level. These findings imply that the use of cut-off levels other than below -2 SDs will not improve the method of population-based SF reference curves.

• Pregnant women are capable of measuring SF heights by themselves. Self-administered SF measurements provide an opportunity to follow fetal growth earlier and more frequently during pregnancy.

• Absolute and relative growth references can be derived from weekly self-administered SF measurements. Relative growth is independent of fetal sex, maternal BMI and parity.

• The *SR method* which is theoretically optimal for detection of altered growth speed was found not to predict increased neonatal morbidity or SGA.

• The *SF method* was better at detecting SGA neonates compared with the *SR method*. SGA neonates, both defined as < -2SD and < the 10th percentile, had a higher morbidity compared with the general study population. To use an SGA definition of < 10th percentile, instead of < -2 SD might be worthwhile.

• Screening and diagnostic methods to improve detection of SGA and IUGR prior to delivery should still be a priority issue.
As antenatal detection of SGA and IUGR pregnancies improves outcome for these neonates, and many of these pregnancies are undetected, a continuous ambition to construct better screening and diagnostic methods is requisite.

Conditional centiles constructed from self-administered SF measurements might improve SF measurements as a screening method and ought to be examined.

Since the SGA definition of < 10th percentile includes a high proportion of infants with neonatal morbidity, a less strict SGA definition than the currently used (< - 2SD) might decrease neonatal morbidity. Studies to determine an optimal SGA definition are needed.

Studies on long term consequences for SGA and IUGR neonates, antenatally detected and not detected, are scarce. Improved knowledge is needed in this field. Here the MBR and the unique personal identification number for all Swedish inhabitants can be of great value in future long term follow up studies. The national health and population registers in the Nordic countries offer excellent opportunities to assess long term effects to exposure during fetal life. Through the personal identification number assigned to each citizen in the Nordic countries and included in national and some regional registers, it is possible to link information from different registers and thereby follow each individual from the beginning of life until death.

Development of an internationally unanimous definition of IUGR would facilitate comparisons between studies and thus improve the total knowledge of IUGR.

Further development of longitudinal models to follow fetal growth will probably facilitate clinical decisions in these pregnancies. Relative growth (growth velocity) although not proven better in this thesis ought to be further examined and tested in different settings. Such studies could focus on high risk pregnancies.
Foster och nyfödda barn som är lågviktiga i förhållande till graviditetslängden (small for gestational age, SGA) har ökad dödlighet och sjuklighet (1). Definitionen för SGA varierar. I Sverige definieras SGA som en födelsevikt < - 2 standardavvikelser (SD) från medelvärdet för graviditetslängd och kön (38), vilket motsvarar c:a 2,5 % av nyfödda barn. Internationellt används oftare definitionen 10e percentilen vilket motsvarar 10 % av de nyfödda. SGA gruppen innehåller både friska (genetiskt små) samt tillväxthämmade barn.

Skillnaden i sjuklighet och dödlighet mot normalstora barn är större om definitionen intrauterin tillväxthämnning (IUGR), används (82). IUGR innebär att fostret inte har uppnått sin tillväxtpotential och indikerar alltid avvikande tillväxt (4). Upp till 90 % av de barn som är SGA vid födseln, växer ikapp sina jämnåriga inom 2 år, vilket indikerar att en stor grupp tillväxthämmade barn finns i SGA gruppen (40).

Om foster med tillväxthämnning identifieras och handläggs på ett adekvat sätt före förlossningen kan dödligheten och sjukligheten reduceras (2, 5).


Under 1980-talet infördes bestämning av graviditetslängd med ultraljud, vilket anses vara det mest korrekt aättet för datering av graviditeter idag (33). Under 1990-talet konstruerades 3 nya svenska referenskurvor för SF måttet, baserade på ultraljudsdatering (8, 18, 19).

Dessa kurvors förmåga att upptäcka SGA graviditeter har ej undersöks. SF metoden är enkel och billig men dess kapacitet att finna lågviktiga barn före förlossningen är till stor del okänd.

För att förbättra kurvornas förmåga att finna tillväxthämmade barn har individuella SF kurvor konstruerats med hänsyn tagen till mammans längd, vikt, paritet, etnicitet samt fostrets kön (45). Fostrets kön, vilket ofta ej är känt, är den enskilt viktigaste faktorn som påverkar dessa referenskurvor.
SF metoden borde kunna förbättras av tätare mätningar samt om mätningarna utförs av en och samma mätare. Dessutom borde upptäckt av IUGR teoretiskt kunna förbättras om den statistiska metoden Shiryaev-Roberts (SR metoden), som upptäcker skillnad i tillväxthastighet istället för absoluta mått, används (27, 69, 70).

Syftet med denna avhandling var att;
1. utvärdera nuvarande metod och referenskurvor för SF mätning.
2. undersöka om självmätning av SF mätt kan utföras.
3. skapa en referenskurva för relativ tillväxt (tillväxthastighet) från självmätta SF mått.
4. utvärdera SR metodens förmåga att finna sjuka och lågviktiga barn före förlossningen.

Studie I
De svenska ultraljudsbaserade referenskurvorna för SF måttet, utvärderades. Studien utfördes som en fall kontroll studie, inkluderande 169 fullgångna och 73 prematurt förlösta SGA barn, samt 296 icke SGA barn som kontroller.

Studien visar att SF metoden har låg sensitivitet för upptäckt av SGA foster. För den kliniskt använda referenskurvan konstruerad av Steingrimsdottir et al (S kurvan) (8) var sensitiviteten 39 % för alla SGA barn, 32 % för fullgångna och 49 % för prematurt förlösta SGA barn. SF kurvan av Kieler et al (61) (K kurvan) hade något högre sensitivitet men något lägre specificitet. När kurvorna analyserades som ROC kurvor visade de sig vara likvärdiga med samma yta under kurvan och utan tydlig optimal brytpunkt för sensitivitet och specificitet.

Studie II
För att få tillgång till tätare SF mätningar, samt mätningar utförda av samma person utfördes en studie med självmätning.

Fyrtio kvinnor med ultraljudsdaterade, singelgraviditeter instruerades av sin barnmorska på mödravårdscentralen att utföra 4 SF mätningar vid varje tillfälle, en gång per vecka från graviditetsvecka 20-25 till förlossning.

Trettiofem kvinnor fullföljde studien. Resultatet visade att självmätningar har högre varians än barnmorskornas, men att det är möjligt att använda självmätning som metod.
Studie III

Studien utfördes för att konstruera en referenskurva för relativ tillväxt (tillväxthastighet) från veckovisa självmätningar av SF måttet.

Sammanlagt inkluderades 300 gravida kvinnor, enligt samma kriterier som i studie II. Mätningar från 191 kvinnor sammanställdes till en absolut och relativ referenskurva för SF-måttet.

Även i denna studie bekräftades att självmätning kan utföras och förefaller möjlig att använda som metod.

För att reducera det individuella inflytandet av moderns längd, vikt, paritet, etnicitet samt fostrets kön konverterades SF måttet till relativ tillväxt, mätt i % jämfört med föregående veckas mätning. Relativ tillväxt av SF måttet utvärderades och föreföll oberoende av de individuella variablerna moderns vikt, paritet och fostrets kön vilket inte var fallet för den absoluta tillväxten.

Studie IV


I studien inkluderades 1888 kvinnor enligt samma kriterier som i studie II och III. Veckovisa självmätningar från 1122 kvinnor användes för utvärdering av SR metoden.

Följande variabler som användes som mått på neonatal sjuklighet; neonatala andningskomplikationer, hypoglykemi, Apgar < 7 vid 5 min, pH < 7.01 i navelartären, en konstruerad samlingsvariabel för ”all neonatal sjuklighet” inkluderande ovan beskrivna variabler samt allvarlig infektion, allvarlig missbildning och neonatal gulsot. Dessutom användes vård vid neonatalavdelning, för tidigt född och operativ förlossning pga. hotande fosterasfyxi.

Resultatet visade att både SR metoden och SF metoden hade låg känslighet för upptäckt av barn med ökad sjuklighet. SF metoden var som förväntat bättre på att upptäcka SGA barn.

SGA barn hade högre sjuklighet än de barn som ej klassificerades som SGA, både för SGA < -2 SD och < 10e percentilen. Dessutom visades att barn som misstänktes vara SGA före förlossningen, med SF metoden, hade ett bättre utfall än barn där denna misstanke ej hade väckts.

Trots teoretiska fördelar med SR metoden kunde den inte visa bättre känslighet för upptäckt av neonatalt sjuka barn vid klinisk användning. Detta kan ha flera förklaringar.
För det första är SR metoden konstruerad för maximal signal i de fall där en totalt avstannad tillväxt inträffar. Detta kanske inte är den vanligaste situationen vid IUGR.

En annan förklaring kan vara att SF mätningar inte är tillräckligt precisa för att på ett bra sätt återge fostrets tillväxt.

Ytterligare förklaring kan vara att förändringar i relativ tillväxt (tillväxthastighet) inte har betydelse förrän fostret har kommit under eller över en viss vikt. Möjligen sker inom detta viktfönster en funktionell adaptation till förändring av tillväxt hos fostret utan ökad sjuklighet som följd.

Slutligen har vi använt neonatal sjuklighet som mått på IUGR, vilket är i linje med rekommendationer för studier om IUGR (81). Dock speglar neonatal sjuklighet en mängd faktorer, förutom IUGR.

Sammanfattningsvis kunde SR metoden i dess nuvarande utformning inte förbättra upptäckten av neonatalt sjuka barn. SGA barn har högre sjuklighet än icke SGA barn. Detta gäller både för SGA < - 2 SD och < 10e percentilen. Barn som misstänks vara SGA före förlossningen har ett bättre utfall än barn med samma födelsevikt, som ej misstänkts vara SGA. En mindre strikt definition av SGA och larmgräns vid screening för SGA kan möjlichen vara till hjälp för att sänka den neonatala sjukligheten. Fler longitudinala studier för ökad förståelse av tillväxthastighet och förbättring av screeningmetoder för upptäckt av SGA behövs.
Acknowledgements

I would like to express my sincere gratitude to

All women who patiently and consistently performed thousands of SF measurements throughout their pregnancies, willing to contribute to improvement of neonatal outcome in the future.

Ove Axelsson, my supervisor, for your infinite patience, your persistent scientific ambition towards improved knowledge. I also appreciate your balance between your professional life and your spare time, with a wide interest in things enrichening life such as; sports, music and reading “The Phantom”. I hope to share many scientific thoughts with you in the future.

Helle Kieler, my co-supervisor, for your stringent attitude to science and medicine, but above all for being such a broad-minded person and generous friend. You are always there to give support, listen and share experiences. Thank you for all the laughs and adventures we have shared since I came to Uppsala.

Max Petzold, my co-supervisor in statistics, for your complete patience with my statistical ignorance and for being such an excellent teacher. All my visits at Nordiska Hälsovårdshögskolan were a pleasure, combined with hard work.

Christian Sonesson, my second statistical advisor, for your crystal clear scientific thoughts and the excellent way you explain them.

Anna Berglund and Ingegerd Lantz, for supporting me with belief in my studies and allowing me to perform my studies in your departments under well organised circumstances.

All midwives at the 27 antenatal clinics participating in the studies, for your positive attitude, recruitment and instructions to almost 2400 pregnant women during the years.

Bo Sultan for providing working conditions enabling me to conduct research and finishing my thesis.
Karin Eurenius, for sharing my room with me for many years and being the best adviser one can hope for both in professional and personal life. You always turn things to the better. We’ve shared many laughs over the years.

Ajlana Lutvica, Peter Lindgren and Solveig Nordén-Lindeberg, for sharing your professional knowledge with me and believing in me. You provide a warm companionship in fetal medicine and obstetrics as well as in all other aspects of life. It is a pleasure to work with you.

Linda Sjögren, my colleague and protégé, for sharing your wisdom of life with me and teaching me about true courage. Your thoughts and our conversations are always rewarding for me. You give me new perspectives of life. You were in my thoughts all through the writing of my thesis.

All my colleagues at the Department of Obstetrics and Gynecology, Uppsala, for sharing one of the best professions in the world with me. We have shared many happy, exciting and sad moments from life together as well as the continuous search for more knowledge.

All personnel at the Antenatal Centre and 96D, for providing the most professional environment I have experienced, for cooperation when giving courses and for your support and encouragement in my research.

My former colleagues in Sundsvall, for introducing me in this profession and giving me all my broad, basic clinical skills and knowledge in obstetrics and gynecology.

My friends, who still are my friends after many years of studies, hard work and family life. I hope to spend more time with you in the future.

In memory of my mother and father, who always encouraged but never forced me to search for knowledge. Who let me think freely not only right and always loved me the way I am.

To my sister and brothers with families, for being my family, appreciating me outside the professional arena, sharing life and precious time with me and my family.

Kicki Moen, for organising my research and lending support while I was busy with clinical work.

Sara, Linnea, Mathilda and Ally, for your ambitious and accurate work at inserting all the data in my studies into the data set.
Owe and Irma, my parents in law, for all the years you were there to support our family.

My treasure; Marcus, Sara and Frida, for giving me the most important thing in my life, love from ones children. You all make me proud of the gifts you have and the way you manage to look out for them. Thank you for putting up with my ambitions.

Gunnar, my beloved husband, for sharing the family life with me, for supporting me in my professional career and bringing a lot of music into my and our children’s life.

This thesis was supported by grants from Uppsala University, Uppsala, Gillbergska stiftelsen, Uppsala, Födelsefonden, Uppsala, Uppsala County Research Council, Uppsala-Örebro Regional Research Council, The Swedish Society of Medicine and Foundation Samariten, Stockholm.
REFERENCES

60. WHO. Principles and Practice of screening for Disease. 1969.
Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 555

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-122256