Beyond the short term effects of caesarean delivery and gynaecological surgery
Kok, N.

Citation for published version (APA):
Kok, N. (2015). Beyond the short term effects of caesarean delivery and gynaecological surgery
Chapter 2

Sonographic measurement of lower uterine segment thickness to predict uterine rupture during a trial of labor in women with previous Cesarean section: a meta-analysis

Ultrasound in Obstetrics and Gynecology 2013

N. Kok
I. C. Wiersma
B. C. Opmeer
I. M. DE Graaf
B. W. Mol
E. Pajkrt
ABSTRACT

Objective
To evaluate the accuracy of antenatal sonographic measurement of lower uterine segment (LUS) thickness in the prediction of risk of uterine rupture during a trial of labor (TOL) in women with a previous Cesarean section (CS).

Methods
PubMed and EMBASE were searched to identify articles published on the subject of sonographic LUS measurement and occurrence of a uterine defect after delivery. Four independent researchers performed identification of papers and data extraction. Selected studies were scored on methodological quality, and sensitivity and specificity of measurement of LUS thickness in the prediction of a uterine defect were calculated. We performed bivariate meta-analysis to estimate summary receiver–operating characteristics (sROC) curves.

Results
We included 21 studies with a total of 2776 analyzed patients. The quality of included studies was good, although comparison was difficult because of heterogeneity. The estimated sROC curves showed that measurement of LUS thickness seems promising in the prediction of occurrence of uterine defects (dehiscence and rupture) in the uterine wall. The pooled sensitivity and specificity of myometrial LUS thickness for cut-offs between 0.6 and 2.0 mm was 0.76 (95% CI, 0.60–0.87) and 0.92 (95% CI, 0.82–0.97); cut-offs between 2.1 and 4.0 mm reached a sensitivity and specificity of 0.94 (95% CI, 0.81–0.98) and 0.64 (95% CI, 0.26–0.90). The pooled sensitivity and specificity of full LUS thickness for cut-offs between 2.0 and 3.0 mm was 0.61 (95% CI, 0.42–0.77) and 0.91 (95% CI, 0.80–0.96); cut-offs between 3.1 and 5.1 mm reached a sensitivity and specificity of 0.96 (95% CI, 0.89–0.98) and 0.63 (95% CI, 0.30–0.87).

Conclusions
This meta-analysis provides support for the use of antenatal LUS measurements in the prediction of a uterine defect during TOL. Clinical applicability should be assessed in prospective observational studies using a standardized method of measurement.
INTRODUCTION
The Cesarean birth rate is rising primarily due to the incidence of elective Cesarean sections (CS) which accounts for one third of CSs. Meanwhile there is a steady decrease in the rate of vaginal birth after Cesarean (VBAC). Combination of the rising number of women with a previous CS and decrease in the VBAC rate suggests an even greater increase in the CS rate in the future. (1)

The VBAC rate has been greatly influenced by clinical studies on the safety of a trial of labor (TOL) after previous CS. Initially a TOL was accepted as safe in the 1980s and early 1990s. However, since the publication of articles questioning the safety of VBAC, there has been a consistent decrease in the VBAC rate. One of the greatest concerns regarding VBAC is the potential for uterine rupture. McMahon et al. in 1996 (2) and Lydon Rochelle et al. in 2001 (3) found that uterine rupture in women with a previous CS was more common after TOL. Nonetheless, the American College of Obstetricians and Gynecologists (ACOG) Committee on Obstetric Practice declared that most women with one previous Cesarean delivery with a low-transverse incision are candidates for and should be counseled regarding VBAC, and should be offered TOL. (4)

The risk of uterine rupture in laboring women with a previous CS varies between 0.2 and 1.5% after induction of labor, compared to 0.5% in women with spontaneous labor after a previous CS. (2,3) On the other hand, there is an increased risk of placenta previa and accreta with every subsequent repeat CS, resulting in higher rates of peripartum hysterectomy. (5) Uterine rupture requires immediate surgical intervention and its occurrence can result in severe morbidity and mortality for infant and mother. (6,7) Accurate prediction of uterine rupture would therefore be extremely valuable, as it would allow women at low risk to proceed with a TOL, whereas women at high risk for uterine rupture could undergo a planned CS.

Several studies have proposed that thinning in the lower uterine segment (LUS) measured by ultrasonography is a predictor of uterine rupture. In 2010 Jastrow et al. conducted a meta-analysis of 12 articles on LUS thickness and risk of uterine scar defect and showed a strong association between the degree of LUS thinning and the risk of uterine defects. (8) However, an ideal LUS thickness cut-off value, usable in clinical practice in women with a scarred uterus, could not be defined. Since the publication of this meta-analysis more data on this topic have become available, justifying a new systematic review. The aim of this study was to identify an optimal LUS thickness cut-off value, defining groups of women with a history of previous CS to whom TOL should either not be offered or could be offered safely.
METHODS

We performed an electronic search of PubMed and EMBASE for relevant articles published between January 1980 and December 2011, using the following keywords: pregnancy, lower uterine segment, Cesarean section, ultrasound and uterine defect (Appendix 1). We searched without language restrictions. We accepted the assistance of colleagues who are part of the Dutch Consortium for Studies in Women’s Health and Reproductivity, who helped translate articles written in their native languages, including Bulgarian, German, Italian and Turkish. We checked the reference lists of primary articles to identify cited articles not captured by electronic searches.

We included studies that reported on pregnant women with at least one previous CS and on the sonographic appearance of the LUS during pregnancy in relation to uterine defects observed during or immediately after delivery. Only studies that allowed construction of two-by-two tables comparing LUS thickness measurement and the occurrence of uterine scar defects, defined as either uterine scar dehiscence or uterine scar rupture, were included. Uterine scar dehiscence was defined as loss of continuity of the myometrial layer without complete rupture of the LUS, also called a uterine ‘window’. Rupture was defined as complete separation of the uterine scar resulting in communication between the uterine and peritoneal cavities. Full LUS thickness was defined as the distance between the bladder wall and the amniotic cavity; this was measured by placing one caliper at the interface between urine and bladder wall and the other at the interface between amniotic fluid and decidual endometrium. Myometrial thickness was defined as the minimum thickness overlying the amniotic cavity at the level of the uterine scar, and in this case only the myometrium was measured.

For each included study, data on study characteristics and test accuracy were independently extracted by four reviewers (B.W.M., E.P., I.C.W., N.K.) using a predesigned data extraction form. Disagreements regarding data extraction were resolved by discussion.

Studies were scored on methodological and clinical characteristics. For methodological quality assessment, we applied the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (9) which uses predefined criteria based on elements of study design, conduct and analysis that relate strongly to bias, variability and quality of reporting of test accuracy. The QUADAS tool, which was developed by a panel of nine experts in the field of diagnostic accuracy, consists of 14 validated questions and does not incorporate a quality score.

The clinical study characteristics extracted included sonographic LUS thickness measurement during pregnancy (index test), definition of uterine scar defect (dehiscence or rupture), full or myometrial LUS measurement, transabdominal (TAS) or transvaginal (TVS) sonographic measurement, level of experience of ultrasound examiners, number of
Sonographic measurement of lower uterine segment thickness

ultrasound examiners, number of measurements, gestational age at LUS measurement, a priori determined threshold for LUS thickness, blinding, setting, study population, study design, data collection, number of participants, adverse neonatal or maternal outcome, VBAC success rate and prevalence of a uterine defect, either dehiscence or rupture.

Statistical analysis

For each study we extracted data to construct a two-by-two table, cross-classifying LUS thickness measured by ultrasound and the presence of LUS defect after delivery. A second reviewer verified all calculations with recalculation. To visualize data we plotted combinations of sensitivity and specificity for each study in the receiver – operating characteristics (ROC) space. A bivariate meta-regression model was used to calculate pooled estimates of sensitivity and specificity and to calculate the corresponding summary ROC (sROC) curve. This method has been described extensively elsewhere. (10 – 12) Briefly, rather than using a single outcome measure per study, such as the diagnostic odds ratio, the bivariate model preserves the two-dimensional nature of diagnostic data expressed as sensitivity and specificity in a single model. The model thereby incorporates the correlation that may exist between sensitivity and specificity within studies due to possible implicit differences in cut-off values (positivity threshold) among studies.

The bivariate model uses a random effects approach for both sensitivity and specificity, allowing for heterogeneity beyond chance due to clinical or methodological differences among studies. In addition, the model acknowledges the difference in precision by which sensitivity and specificity have been measured in each study. This means that studies with a larger number of patients with a uterine defect receive more weight in the calculation of the pooled estimate of sensitivity, while studies with more patients without a uterine defect are more influential in the pooling of specificity.

Different studies reported accuracy based on different cut-off values, and some studies also reported accuracy for more than one cut-off. With the available data, separate sROC-curves were calculated for full and myometrial LUS measurements. The bivariate model was fitted using data for all reported cut-offs in each case, thus reflecting the change in accuracy associated with a shift in threshold. In addition, sROC curves were estimated based on the reported accuracy for a limited range of cut-offs (2.0 – 3.0 mm and 3.1 – 5.1 mm for full thickness and 0.6 – 2.0 mm and 2.1 – 4.0 mm for myometrial thickness), reflecting pooled estimates of sensitivity and specificity for each threshold range. Bivariate models were fitted using Proc NLMixed (SAS 9.3 for Windows; SAS Institute Inc, Cary, NC, USA).
Table 1. Characteristics of studies assessing the association between lower uterine segment (LUS) measurement and occurrence of uterine defect during trial of labor in women with a previous Cesarean section (CS)

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Study duration (years)</th>
<th>Study design</th>
<th>Blinding</th>
<th>Participants with analyzable data (n)</th>
<th>Sono-</th>
<th>LUS measurement</th>
<th>Reported outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valclavinkova (1984)33</td>
<td>Sweden</td>
<td>03</td>
<td>Prosp.</td>
<td>NR</td>
<td>68</td>
<td>NR</td>
<td>Myometrium</td>
<td>Thinning</td>
</tr>
<tr>
<td>Michaels (1988)16</td>
<td>USA</td>
<td>30</td>
<td>Prosp.</td>
<td>None</td>
<td>58</td>
<td>TAS</td>
<td>Full</td>
<td>Thinning</td>
</tr>
<tr>
<td>Fukuda (1988)22</td>
<td>Japan</td>
<td>47</td>
<td>NR</td>
<td>NR</td>
<td>84</td>
<td>TAS</td>
<td>Myometrium</td>
<td>Thinning</td>
</tr>
<tr>
<td>Fukuda (1991)15</td>
<td>Japan</td>
<td>54</td>
<td>NR</td>
<td>NR</td>
<td>216</td>
<td>TAS/TVS</td>
<td>Myometrium</td>
<td>Thinning</td>
</tr>
<tr>
<td>Popov (1994)27</td>
<td>Bulgaria</td>
<td>29</td>
<td>Retro.</td>
<td>None</td>
<td>26</td>
<td>TAS</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Hebsch (1994)23</td>
<td>Germany</td>
<td>NR</td>
<td>Prosp.</td>
<td>Single</td>
<td>10</td>
<td>TAS/TVS</td>
<td>Myometrium</td>
<td>Thinning</td>
</tr>
<tr>
<td>Tanik (1996)32</td>
<td>Turkey</td>
<td>04</td>
<td>NR</td>
<td>NR</td>
<td>50</td>
<td>TAS</td>
<td>Full</td>
<td>Thinning</td>
</tr>
<tr>
<td>Rozenberg (1996)14</td>
<td>France</td>
<td>57</td>
<td>Prosp.</td>
<td>NR</td>
<td>642</td>
<td>TAS</td>
<td>Full</td>
<td>Dehiscence/Rupture</td>
</tr>
<tr>
<td>Qureshi (1997)28</td>
<td>Japan</td>
<td>19</td>
<td>Prosp.</td>
<td>NR</td>
<td>43</td>
<td>TVS</td>
<td>Myometrium</td>
<td>Dehiscence/Rupture</td>
</tr>
<tr>
<td>Montanari (1999)25</td>
<td>Italy</td>
<td>28</td>
<td>Prosp.</td>
<td>None</td>
<td>61</td>
<td>TVS</td>
<td>Myometrium</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Rozenberg (1999)29</td>
<td>France</td>
<td>19</td>
<td>Prosp.</td>
<td>None</td>
<td>198</td>
<td>TAS</td>
<td>Full</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Suzuki (2000)31</td>
<td>Japan</td>
<td>30</td>
<td>NR</td>
<td>NR</td>
<td>83</td>
<td>TAS</td>
<td>Myometrium</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Asakura (2000)37</td>
<td>Japan</td>
<td>41</td>
<td>NR</td>
<td>Single</td>
<td>186</td>
<td>TVS</td>
<td>Myometrium</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Cheung (2005)120</td>
<td>Canada</td>
<td>19</td>
<td>NR</td>
<td>None</td>
<td>102</td>
<td>TVS/TAS</td>
<td>Myometrium</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Mohammed (2010)26</td>
<td>Egypt</td>
<td>07</td>
<td>Prosp.</td>
<td>None</td>
<td>100</td>
<td>TAS/TVS</td>
<td>Myometrium</td>
<td>Dehiscence</td>
</tr>
<tr>
<td>Kushtagi (2011)24</td>
<td>India</td>
<td>NR</td>
<td>Prosp.</td>
<td>None</td>
<td>106</td>
<td>TAS</td>
<td>Full</td>
<td>Dehiscence</td>
</tr>
</tbody>
</table>

Only the first author of each study is given. All studies included women with at least one previous CS, except that by Popov which included cases with only one previous CS. NR, not reported; Prosp., prospective cohort; Retro., retrospective cohort; TAS, transabdominal sonography; TVS, transvaginal sonography.
RESULTS
Figure 1 summarizes results of the literature search. The initial search identified 147 citations, of which 113 were excluded after screening of titles and abstracts. In the 34 studies that were selected for further reading, 10 were excluded because of inappropriate reporting of outcome and three were excluded because of inaccessibility (i.e. published in languages that none of our colleagues could translate). Finally we included 21 studies (13–33) reporting on 2776 women with analyzable data concerning LUS measurements and uterine defects.

Table 1 shows characteristics of the included studies, all of which were cohort studies. There were 14 prospective studies and one retrospective study, and six studies did not report whether they were prospective or retrospective. Sample sizes ranged from 10 to 642 women (median 71; interquartile range, 149.5). Only five studies were blinded, i.e. sonographic findings were not conveyed to the treating obstetrician and decisions for repeat CS were performed because of obstetric indications only.

Only nine studies reported fetal outcome, of which eight studies found no adverse outcome, and in one study two infants died. Eleven studies that reported on maternal outcome showed no maternal complications. The remaining 10 articles did not report adverse maternal outcome, apart from CS.

Gestational age at LUS measurement ranged from 34 to 39 weeks. Nine studies reported the number of ultrasound examiners, and only two studies described the level of experience of the ultrasound examiner. Myometrial thickness was measured in 13 studies; full LUS thickness was measured in six studies; and in one study both myometrial and full LUS thickness were measured. One study did not specify the method of measurement. TAS was used in nine studies, TVS in five and both techniques were used in six studies.
Figure 2. Summary of results of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool for articles included in the present analysis (yes, no, unclear).

Figure 3. Estimated summary receiver-operating characteristics curves (—) and pooled sensitivity and specificity (*) of: (a) myometrial lower uterine segment (LUS) thickness measurement and (b) full LUS thickness measurement, for the prediction of uterine defects during a trial of labor in women with a previous Cesarean section. Rectangles show the observed accuracy for each cut-off point reported in each study.
The cut-off values used to define an insufficient LUS ranged from 0.5 to 3.0 mm in studies measuring myometrial LUS thickness and from 1.5 to 5.1 mm in those using full LUS thickness. Twelve articles defined their cut-off value a priori. With respect to definition of the outcome, six studies reported on thinning and dehiscence of the LUS, nine on dehiscence, four on the combination of dehiscence and rupture and one on incomplete rupture. One study did not provide a clear definition of LUS defect.

Figure 2 summarizes results of quality assessment of the studies. The majority of studies gave clear descriptions of selection criteria, methods of measurements and reference standards. The spectrum of patients and the clinical data used in most studies were representative of common practice. The weakness of most studies was the absence of reporting on uninterpretable test results and of blinding, making it difficult to assess bias.

**Data analysis**

The VBAC rate (\(\text{VBAC rate} = \frac{\text{number of VBAC}}{\text{number of all women with a previous CS}} \times 100\)) varied from 19 to 68% in the included studies. The VBAC success rate (\(\text{VBAC success rate} = \frac{\text{number of VBACs}}{\text{number of women undergoing a TOL}} \times 100\)) varied from 20 to 78%.

Figure 3 shows sROC curves for the capacity of myometrial and full LUS thickness measurement in the prediction of uterine defects at different cut-off values. The sROC curve reflects the change in accuracy associated with a shift in threshold.

Figure 4 shows sROC curves for the capacity of full LUS thickness measurement in the prediction of uterine defect for cut-off ranges of 2.0 – 3.0 mm and 3.1 – 5.1 mm. Full LUS thickness measurement using cut-offs between 2.0 and 3.0 mm reached a specificity of 0.91 (95% CI, 0.80 – 0.96) at a sensitivity of 0.61 (95% CI, 0.42 – 0.77). Full LUS thickness measurement using cut-offs between 3.1 and 5.1 mm reached a specificity of 0.63 (95% CI, 0.30–0.87) at a sensitivity of 0.96 (95% CI, 0.89–0.98).

The accuracy of TVS and TAS could not be compared statistically because of the limited number of studies. We therefore cannot conclude whether TVS has a better capacity for predicting uterine defects than does TAS. Ideally, sufficient data would be available for stratification of the use of TAS and TVS and each cut-off, so that pooled estimates could be obtained for each combination.

**DISCUSSION**

In this Review we describe the value of LUS thickness measurement in the prediction of uterine defects, either complete rupture or dehiscence at birth, during TOL in women who had undergone a previous CS. The most important finding is the strong negative correlation between LUS thickness and risk of uterine defect. Shapes of the estimated sROC curves for myometrial and full LUS thickness were very similar, indicating no
significant difference in any of the three parameters: accuracy, shape and position. Thus, in our analysis of pooled data, these two methods were found to be equivalent.

We found that a full LUS thickness cut-off of 3.1–5.1 mm and a myometrium thickness cut-off of 2.1 – 4.0 mm provided a strong negative predictive value for the occurrence of a defect during TOL. Full LUS thickness cut-off of 2.0 - 3.0 mm and a myometrium thickness cut-off between 0.6 and 2.0 mm provided a strong positive predictive value for the occurrence of a defect.

The strength of this review is our thorough search without language restrictions. We were able to include 21 articles that described LUS thickness in relation to the occurrence of a uterine defect during delivery. In 2010, Jastrow et al. conducted a meta-analysis of 12 articles on LUS thickness and risk of uterine scar defect but could not define an ideal LUS thickness cut-off value usable in clinical practice (8). Although we found an additional nine articles, unfortunately this did not enable us to determine an ideal cut-off value. We would have liked to establish the accuracy of LUS thickness at the same cut-offs for all studies; however, we had to analyze pooled data because of different cut-offs in the different studies. This variation in LUS thickness cut-off values reflects the considerable amount of heterogeneity among studies and is one of the limitations of our study. Articles were retrieved if they allowed construction of two-by-two tables comparing LUS thickness

Figure 4. Plots of estimated summary receiver-operating characteristics (sROC) curves and pooled sensitivity and specificity (sens/spec) for the capacity of: (a) myometrial lower uterine segment (LUS) thickness measurement for cut-off ranges 0.6-2.0 mm (○, observed accuracy; •, pooled sens/spec; —, sROC curve) and 2.1-4.0 mm (□, observed accuracy; ■, pooled sens/spec; --, sROC curve) and (b) full LUS thickness measurement for cut-off ranges 2.0-3.0 mm (○, observed accuracy; •, pooled sens/spec; —, sROC curve) and 3.1-5.1 mm (□, observed accuracy; ■, pooled sens/spec; --, sROC curve), in the prediction of uterine defects during a trial of labor in women with a previous Cesarean section, ----, 95% confidence regions for estimated pooled sens/spec.
and occurrence of uterine defect. Fewer than half of the studies defined an a priori cut-off value at the start of their study. Some investigators opted for the optimal cut-off, considering their own data, resulting in overly optimistic estimates of prognostic accuracy. By analyzing data using a bivariate meta-regression model we tried to incorporate the correlation between sensitivity and specificity due to implicit differences in cut-off value (positivity threshold) among studies.

Another concern in this area of research is the large number of relatively small studies that we identified. Small studies are inclined to overestimate the predictive capacity of LUS thickness in the prediction of a uterine defect, whereas larger studies generally tend to report less extreme results. (34) We tried to minimize any overestimation of predictive capacity by using the bivariate meta-regression model to calculate weighted pooled estimates of sensitivity and specificity according to differences in sample size and precision of studies. However, we could not correct for the fact that more than 75% of studies were not blinded, indicating that the doctors performing delivery were aware of the measured LUS thickness and decisions regarding repeat CS could have been based on these findings, alongside obstetric indications.

A further limiting factor was the variable definition of uterine defect among studies, ranging from thinning to complete rupture. Uterine dehiscence is known to be asymptomatic and the absence of clinical significance of ‘silent’ scar dehiscence has been mentioned by Peaceman and Sciarra (35) and by Petrikovsky. (36) Dehiscence was the most frequently observed primary outcome in the majority of studies. Although thinning, dehiscence and possibly rupture are likely to represent different expressions of a similar process, the real challenge in clinical practice lies in the accurate prediction of uterine rupture. Since this event is rare, only five of the included studies actually reported on LUS thickness as a predictor of uterine rupture. Another important limitation is the fact that the measurement of LUS thickness has not yet been standardized. There is no consensus among studies regarding which layer(s) of the LUS should be measured or by which route. Both TAS and TVS, as well as combinations of these, have been used. Sen et al. showed the correlation between TAS and TVS in measuring LUS thickness to be excellent (30). However, a more recent study showed interobserver agreement to be better when TVS was used. (37)

Identifying women at risk for uterine rupture remains an important challenge in obstetric care. Before the measurement of LUS thickness can be implemented in clinical practice, analytical and clinical validity and utility must be addressed. At present, we are only at the beginning of this process. An ideal screening test to predict uterine rupture would require high levels of both sensitivity and specificity (≥90%). If such a test were to become available, it is very likely that this would influence medical decision-making, i.e. through the accurate selection of women with a scarred uterus unlikely to have uterine rupture and
therefore suitable for a TOL, as opposed to women with a scarred uterus likely to have a uterine rupture and therefore suitable for repeat CS. Ultimately, it is the influence of implementation of a test on patient outcomes that matters.

We found a median VBAC success rate of 54%. This rate corresponds with the present literature, in which the reported rate of successful TOL varies from 43% to 80%, increasing to almost 90% after a previous vaginal birth, including VBAC. (38–40) Unfortunately, no studies have compared LUS thickness between women with and without previous vaginal births who have had a previous CS, and the appropriate cut-off value to be used in women with previous vaginal birth or VBAC needs to be determined. Moreover, other factors such as a short interval between deliveries, more than one previous CS, prior classical hysterotomy, treatment with uterotonic agents, postpartum fever and maternal age may influence the accuracy of this tool and should be further investigated. Consequently, there is a strong need for large cohort studies in which the result of LUS measurements is not disclosed to attending physicians until after the delivery.

ACKNOWLEDGEMENTS
We are grateful for the assistance of colleagues, part of the Dutch Consortium for Studies in Women’s Health.
REFERENCES


