Prenatal diagnosis of LUTO: how to improve diagnostic accuracy

Short title: Prenatal diagnosis of LUTO

F. Fontanella¹, L. K. Duin¹, P. N. Adama van Scheltema², T. E. Cohen – Overbeek³, E. Pajkrt⁴, M. Bekker⁵–⁶, C. Willekes⁷, C. J. Bax⁸, V. Gracchi⁹, D. Oepkes² and C.M. Bilardo¹

¹Department of Obstetrics, Gynaecology and Prenatal Diagnosis, University Medical Center Groningen, University of Groningen, The Netherlands;
²Department of Obstetrics, Gynaecology and Prenatal Diagnosis, Leiden University Medical Center, Leiden, The Netherlands;
³Department of Obstetrics and Gynaecology, Division of Obstetrics and Prenatal Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands;
⁴Department of Obstetrics, Academic Medical Center Amsterdam, Amsterdam, The Netherlands;
⁵Department of Obstetrics, Gynaecology and Prenatal Diagnosis, Radboud University Medical Center, Nijmegen, The Netherlands;
⁶Department of Obstetrics, Gynaecology and Prenatal Diagnosis, University Medical Center Utrecht, Utrecht, The Netherlands;
⁷Department of Obstetrics, Gynaecology and Prenatal Diagnosis, University Medical Center, Grow School for Oncology and Medical Biology, Maastricht, The Netherlands.
⁸Department of Obstetrics, Gynaecology and Prenatal Diagnosis, VU University Medical Center, Amsterdam, The Netherlands.
⁹Department of Pediatrics, University Medical Center Groningen, University of Groningen, The Netherlands.

Corresponding author:
F.Fontanella

Address: Hanzeplein 1, Groningen

e-mail: federica.fontanella@gmail.com; f.fontanella@umcg.nl

Keywords: LUTO, Lower urinary tract obstruction, Prenatal diagnosis, Megacystis

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.18990
**Objective:** to propose a clinical score for the optimal antenatal diagnosis of Lower Urinary Tract Obstruction (LUTO) in the second trimester of pregnancy, as alternative to the commonly used ultrasound (US) triad (megacystis, keyhole sign and hydronephrosis).

**Methods:** This was a national retrospective study carried out at the eight tertiary Fetal Medicine Units (FMUs) in the Netherlands. Only cases referred for megacystis starting from the second trimester and with clear postnatal diagnosis were included in the study. The following antenatal data was collected at referral: amniotic fluid volume, renal cortical appearance, bladder volume, fetal ascites, ureteral size, keyhole sign, fetal sex and gestational age at referral. A multivariate analysis was performed starting with the inclusion of all antenatal variables and then excluding the weakest predictors by the backward stepwise strategy.

**Results:** Over a seven-year period, 312 fetuses with a diagnosis of megacystis were referred to one of the eight Dutch tertiary FMUs. A final diagnosis was achieved in 143 cases, including 124 cases of LUTO and 19 cases reclassified after birth as non-obstructive megacystis. The optimal bladder volume cut-off for prediction of LUTO was 35 cm$^3$ (Area Under Curve (AUC)=0.7, $p=0.03$). On the base of the multivariate analysis, a clinical score was formulated. This included: fetal sex, degree of bladder distension, ureteral size, presence of oligohydramnios and GA at referral. The combination of these five variables demonstrated good accuracy in discriminating LUTO from non-obstructive megacystis (AUC=0.84), compared to the poor performance of the US triad (AUC=0.63, $p=0.07$).

**Conclusions:** We propose a clinical score that combines five antenatal variables for the prospective diagnosis of congenital LUTO. This score showed good discriminative capacity in predicting LUTO, and better diagnostic accuracy compared to the classic US triad. Future studies to validate these results are needed in order to refine the antenatal management of LUTO and prevent inappropriate fetal interventions.
Introduction

The term Lower Urinary Tract Obstruction (LUTO) refers to an heterogeneous group of anatomical anomalies causing an obstruction in the urethra (1). During fetal life, LUTO entails a sequence of events detectable at the antenatal US scan. This typically starts with the evidence of distended fetal bladder (megacystis) accompanied by hydronephrosis and progressing to renal dysplasia with abnormal renal parenchymal appearance at US scan and, eventually resulting in severe oligohydramnios (2). The condition has high mortality and postnatal morbidity due to lung hypoplasia and impaired renal function (3). When LUTO is suspected in the first trimester and a megacystis greater than 12 mm is seen, the prognosis is extremely poor and parents often opt for termination of pregnancy (4)(5). For cases identified later in pregnancy no hard criteria for defining LUTO and predicting the exact prognosis have yet been proposed (6). Beyond the first trimester, the diagnosis of LUTO is typically based on the evidence of three US findings: megacystis, dilated posterior urethra (known as keyhole sign), and either unilateral or bilateral hydronephrosis.

Over the last 20 years, fetal therapy has been attempted on the assumption that by relieving the intra-cavitary pressure caused by the obstruction, mortality ad renal damage could possibly be prevented. The PLUTO trial investigated this assumption, demonstrated a significant improvement of survival in fetuses treated with vesico-amniotic shunt, and reported an high morbidity among survivors irrespective of the antenatal management (7). To date, whether and when the in-utero treatment should be offered remains a matter of debate, and the eventual selection of candidates is still suboptimal, owing to the high number of false-positive LUTO cases. In fact, a previous study has reported that one third of all prenatally suspected LUTO cases are reclassified postnatally, the majority of them to vesico-ureteral reflux (VUR) (9). For this reason, an improvement in the diagnostic accuracy of US for the diagnosis of LUTO is needed.

The aim of this study was to identify the optimal combination of US parameters for the antenatal diagnosis of LUTO as alternative to the commonly used LUTO triad (megacystis, keyhole sign and hydronephrosis).
Methods

This was a retrospective national study carried out at all eight Fetal Medicine Units (FMUs) of University Hospitals in the Netherlands. Cases were collected according to the start of registration in databases: this was from 2000 to 2015 in three centers (Erasmus Medical Center, Rotterdam; Academic Medical Center, Amsterdam; University Medical Center, Maastricht); from 2004 to 2015 in two centers (University Medical Center Groningen and Radboud University Medical Center, Nijmegen); and between 2007 and 2014 in the remaining centers (Leiden University Medical Center, Leiden; Utrecht University Medical Center, Utrecht; VU University Medical Center, Amsterdam). These FMUs act as expert referral center for all anomalies suspected in peripheral hospitals and external ultrasound clinics in the Netherlands. We only included cases referred for fetal megacystis diagnosed from the 18th week’s gestation onwards, thus either directly after the routine second-trimester scan or after US scans performed later in pregnancy for growth or other obstetrical indications. Fetal megacystis was defined as an enlarged bladder failing to empty during an extended US examination lasting at least 40 minutes.

The following antenatal data were collected at referral: gestational age, fetal sex, evidence of keyhole sign, or fetal ascites (caused by leakage or rupture of the distended bladder), hydronephrosis, amount of amniotic fluid, renal cortical appearance, size of right and left ureteral diameter, and antero-posterior, transverse and longitudinal bladder diameter. The bladder volume was calculated using the formula: longitudinal diameter x transverse diameter x antero-posterior diameter x π/6 (10). The sum of right and left ureteral diameter was calculated for each case, and for ureters non-visualized at the US scan the sum was considered as zero millimeters. Amniotic fluid was considered reduced in case of a single deepest pocket (SDP) of less than 2 cm. US reports were reviewed in order to retrieve data on amniotic fluid volume, renal pelvis dilatation and renal parenchymal echogenicity. Only the first detailed US report at referral was used for analysis.

Outcome data included all available postnatal data on surgeries and medical examinations for live-born infants, and postmortem examinations for perinatal deaths, when available. The term
LUTO referred to a group of anatomical anomalies causing urethral obstruction (1). This group thus included cases with posterior urethral valves (PUV), urethral stenosis, urethral atresia, and also cases with LUTO reported as final diagnosis, but without further details concerning the type of obstruction (non-specified LUTO).

Antenatal baseline characteristics were compared using the chi-square test or Fisher’s exact test for categorical variables and the Student’s t test for continuous variables. Univariate analysis was performed to examine the association between candidate predictors and final diagnosis. A logistic model was developed considering eight variables first, and then proceeding by the backward stepwise strategy in excluding progressively the weakest predictors. The model performances were assessed by the Hosmer-Lemeshow test for goodness of fit, and the discriminative performance in the model was evaluated by the area under the receiver-operating characteristic (ROC) curve using the predicted and the actual outcome. The model was internally validated with bootstrapping using R-project software 3.4.2 (https://www.r-project.org/; package rms).

A clinical score was developed based on the results of the logistic model. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated. Data analyses were performed using the statistical software package SPSS Statistics 23 (SPSS Inc., Chicago, IL, USA).
Results

During the study period in total 312 pregnancies were referred because of a suspected fetal megacystis from the 18th week onwards. The outcome of the 312 pregnancies was: 71 terminations of pregnancies (TOP), 10 cases of intrauterine fetal death (IUFD), 38 neonatal deaths, and 193 live-born infants. Ninety-eight cases (31%) were excluded from the study because of missing or incomplete data not enabling a final diagnosis (in 55 cases autopsy was declined and 43 cases were lost to follow-up) and 68 cases (22%) were excluded because of incomplete antenatal data or measurements. Moreover, three megacystis cases not suspected for LUTO were excluded from further analysis. They all presented polyhydramnios and macrosomia and an overgrowth syndrome was confirmed after birth (Figure 1).

Based on postnatal investigations or postmortem examinations, a final diagnosis was achieved in 143 cases, including 124 (87%) true LUTO cases (74 PUV, 4 urethral atresia, 6 urethral stenosis and 40 non-specified LUTO) and 19 (13%) cases reclassified postnatally as non-obstructive megacystis (12 infants with VUR (vesico-ureteral reflux), 4 cases of primary mega-ureters, one fetus with megacystis-microcolon-intestinal hypoperistalsis syndrome, and 2 cases without any evidence of urological anomaly and normal voiding at birth).

Descriptive statistics, sensitivity, specificity and details on the univariate analysis according to final diagnosis, are presented in Table 1 and 2. Longitudinal bladder diameter (LBD) showed poorer accuracy at the univariate analysis compared to bladder volume, and was therefore excluded from further analysis. A ROC curve of bladder volume was performed to identify the optimal cut-off for prediction of LUTO (AUC = 0.66 (CI 0.6-0.8), p = 0.03, optimal cut off: 35 cm$^3$). In addition, 13 cases showed urinary ascites at referral with a collapsed urinary bladder, suggesting that the a severely enlarged bladder had ruptured and thus making bladder volume measurement no longer reliable. Therefore, severe megacystis was defined by a bladder volume > 35 cm$^3$ or ascites at referral.

A multivariate logistic analysis was performed starting with the inclusion of all the following antenatal variables considered as theoretically relevant in the literature for the prospective
diagnosis of LUTO: renal cortical appearance (normal or abnormal), amniotic fluid volume (normal or reduced), gestational age at referral (before or after 28 weeks' gestation), degree of bladder distension (mild or severe), fetal sex (female or male), evidence of keyhole sign (yes or no), and ureteral size, as continuous variable. The stepwise backward method resulted in the progressive collapse of the variables with poorer performance. These were: fetal hydronephrosis, renal cortical appearance and keyhole sign.

The final model included five predictive variables: severe megacystis (odds ratio (OR): 4.21 (95% CI, 0.98-18.21), p = 0.05; after bootstrapping: p = 0.06, 95% CI: 0.1-20.1), ureteral size (OR: 1.25 (95% CI, 1.02-1.54), p = 0.04; after bootstrapping: p = 0.02, 95% CI: 0.1-0.8), oligohydramnios (OR: 3.7 (95% CI, 0.71-19.25); p = 0.12; after bootstrapping: p = 0.08, 95% CI: -0.4-19.9), male sex (OR: 4.09 (95% CI, 0.88-18.88); p = 0.07; after bootstrapping: p = 0.03, 95% CI: -0.1-4.2) and referral before the 28th week (OR: 3.72 (95% CI, 1.18-11.72; after bootstrapping: p = 0.15, 95% CI: 0.2-2.8); p = 0.025). The Hosmer-Lemeshow test for goodness of fit showed a good fit of this model with p = 0.94, considering that p-values closer to one indicates a better fit. The optimism-corrected model performance after bootstrapping was 82%, thus 2% smaller than with the original dataset.

A clinical score was formulated based on the results of the logistic model (Table 3). Figure 2 shows the accuracy in discriminating LUTO from non-obstructive megacystis of this proposed clinical score compared to a theoretical model only based on the commonly used LUTO triad (Area Under Curve: 0.84 (95% CI, 0.75-0.93), p < 0.001) vs 0.63 (95% CI, 0.49-0.77), p = 0.08). A ROC curve analysis identified at 9.5 the optimal cut-off point for the clinical LUTO score in predicting the risk of LUTO (Table 3; sensitivity 78%, specificity 79%).
Discussion

In this study, we propose a clinical score for calculating the risk of congenital LUTO during pregnancy, based on five antenatal variables all evaluated at the detailed US scan at referral: bladder distension (severe or moderate), bilateral ureteral dilatation (as continuous variable), amount of amniotic fluid (normal or oligohydramnios), fetal sex and GA at referral (before or after the 28th week). This score demonstrated good discriminative value in distinguishing real LUTO from non-obstructive megacystis, which would not be amenable for antenatal treatment, and a better performance than the classic antenatal triad. The use of this new combination of US parameters enables an optimal identification of LUTO cases at the time of referral, allowing for appropriate counseling and management options.

The role of fetal therapy for LUTO is still debated in the literature and the opportunity to gain high-quality evidence has been missed with the premature conclusion of the PLUTO trial. A retrospective multicenter study was recently published with the aim to explore the effectiveness of fetal therapy in cases with severe LUTO, defined by megacystis, increased bladder-wall thickness, bilateral severe hydronephrosis and oligohydramnios. Despite the strict criteria, still 23% of treated fetuses were wrongly suspected for LUTO (6). We think that both disease severity and selection of candidates for in-utero treatment are influential determinants for fetal therapy's effectiveness, and that an improvement in the diagnostic accuracy of antenatal US is thus needed.

In this study, fetal hydronephrosis was observed in 88% of LUTO cases, and in 79% of non-obstructive megacystis cases. A recent review reported hydronephrosis even only in 40-50% of LUTO cases and questioned the strength of this association (11). The keyhole sign demonstrated high specificity (74%), but poor sensitivity for LUTO (48%; Table 2). Other studies (12)(13)(14) have previously reported poor accuracy of this US sign for the prospective diagnosis of LUTO and in particular PUV. It has been hypothesized that a possible explanation for its low reliability is that miscellaneous types of bladder dysfunctions such as detrusor instability and bladder-sphincter dyssynergy can cause a dilatation of the bladder neck. The latter has been in fact diagnosed in 30% of male infants with VUR at voiding cystourethrograms.
The dilatation of the bladder neck at the prenatal ultrasound scan could mimic a dilated posterior urethra, with an ultrasound appearance similar to a keyhole, without being a true dilatation of the posterior urethra (13)(16). Therefore, although keyhole sign and hydronephrosis have been considered thus far as key findings of LUTO, and PUV in particular, they are poorly predictive of the exact postnatal diagnosis.

The amount of amniotic fluid was included in the final model, although it demonstrated poor sensitivity at the univariate analysis. In our cohort in fact 69/124 (56%) of LUTO cases showed normal amniotic fluid at referral. This is consistent with previous studies that reported a rate of 39% (13), albeit they did not set a specific gestational age for evaluating this parameter. Oligohydramnios is typically considered as a characteristic finding of LUTO, as cases with non-obstructive megacystis are unlikely to develop such anomaly. For this reason, a reduced amount of amniotic fluid is often used as an eligibility criterion for the in-utero treatment. However, some fetal medicine expert argue that oligohydramnios occurs when renal parenchyma has already been severely damaged and for this reason in-utero treatment should rather be offered to candidates with normal or just moderately reduced amniotic fluid volume. Our results confirm that the amount of amniotic fluid, evaluated at the first detailed US scan, can be used in evaluating the risk of LUTO. However its absence should not rule out the antenatal suspicion of LUTO.

Another interesting finding of this study was that the severity of bladder distension is an independent predictor of LUTO. Our results showed that the likelihood of LUTO was four folds higher (OR 4.21, p = 0.05) in case of severe megacystis, defined by bladder volume greater than > 35 cm³ or fetal ascites at referral. Megacystis has always been considered a key feature of LUTO with a high predictive value (13). However, an objective threshold to define megacystis during the second and third trimester is still lacking. Previous studies (10)(17) have investigated fetal urine production according to GA in healthy fetuses without reporting fetal bladder dimensions. Miscellaneous criteria have been used in the literature for defining megacystis in the second and third trimester, ranging from a bladder length >99th percentile for gestational age in absence of a normogram (18), to the most commonly used definition of a fetal bladder failing
to empty during a period of 45 minutes (19)(20)(21)(22). To sum up, at present there have not yet been published prospective studies that have elucidated the normal bladder dimensions during the second and third trimester of pregnancy and defined a threshold for pathological enlargements. This is therefore urgently needed.

A strength of this study is that, at variance with previous studies, the antenatal variables were combined in a multivariate analysis and were only considered at the moment of referral. In spite of the fact that this approach has reduced our cohort to 143 cases, this is thus far the largest study in the literature.

Some critical remarks need to be made. First, not all the antenatal variables were directly measured by the sonographer, and in a proportion of cases the first author (FF) measured them retrospectively on stored pictures. This also implied that serial bladder measurements were not available in most of the cases. Second, the retrospective design implied that for the renal parenchymal appearance we had to rely upon the subjective judgment reported on the US reports, rather than on an objective measurement. This limitation may have affected the accuracy of this variable.

Statistical bootstrapping showed that the overfitting of the model was small, suggesting that the model could hold for the overall population suspected for LUTO. However, an external validation is essential before endorsing this statement and supporting its clinical applicability.

To conclude, in order to improve the diagnostic accuracy of LUTO in the second and third trimester, the criteria that need to be evaluated are: fetal bladder enlargement, ureteral dilatation, gestational age at referral, fetal sex and evidence of oligohydramnios. Future studies for validating externally these results are needed in order to refine the antenatal identification of LUTO, prevent unnecessary fetal interventions and optimize the prenatal management.
References


Table 1. Antenatal ultrasound characteristics according to the final diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>Non-obstructive megacystis (n = 19)</th>
<th>LUTO (n = 124)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex (male)</strong></td>
<td>14 (74%)</td>
<td>115 (93%)</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Keyhole sign</strong></td>
<td>5 (26%)</td>
<td>59 (47%)</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Echogenic kidneys</strong></td>
<td>4 (21%)</td>
<td>67 (54%)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><strong>Oligo- or anhydramnios</strong></td>
<td>2 (11%)</td>
<td>55 (44%)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><strong>Hydronephrosis</strong></td>
<td>15 (79%)</td>
<td>108 (88%)</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>Referral before the 28th weeks</strong></td>
<td>8 (42%)</td>
<td>82 (66%)</td>
<td>0.04</td>
</tr>
<tr>
<td>(median and range, weeks)</td>
<td>25 (19 – 36)</td>
<td>23 (18 – 36)</td>
<td></td>
</tr>
<tr>
<td><strong>Bladder volume (cm³)</strong></td>
<td>19.5</td>
<td>61.2</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>(median and range)</td>
<td>18 (0.7 – 58)</td>
<td>31 (0.3 – 390)</td>
<td></td>
</tr>
<tr>
<td><strong>Bladder longitudinal diameter (mm)</strong></td>
<td>13</td>
<td>17</td>
<td>0.02</td>
</tr>
<tr>
<td>(median and range)</td>
<td>38 (15 – 55)</td>
<td>43 (9 – 91)</td>
<td></td>
</tr>
<tr>
<td><strong>Ureteral size (mm)</strong></td>
<td>1.1</td>
<td>5.7</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>(sum of right and left ureteral diameters)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Univariate analysis, sensitivity and specificity with regard to final diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>OR (95% CI)</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male)</td>
<td>4.56 (1.3 – 15.6)</td>
<td>93%</td>
<td>26%</td>
</tr>
<tr>
<td>Keyhole sign</td>
<td>2.54 (0.8 – 7.5)</td>
<td>48%</td>
<td>74%</td>
</tr>
<tr>
<td>Echogenic kidneys</td>
<td>4.41 (1.4 – 14.0)</td>
<td>54%</td>
<td>79%</td>
</tr>
<tr>
<td>Oligo- or anhydramios</td>
<td>6.78 (1.5 – 30.6)</td>
<td>44%</td>
<td>90%</td>
</tr>
<tr>
<td>Referral before the 28th week</td>
<td>2.69 (1.0 – 7.2)</td>
<td>66%</td>
<td>58%</td>
</tr>
<tr>
<td>Hydronephrosis</td>
<td>1.92 (0.6 – 6.6)</td>
<td>88%</td>
<td>21%</td>
</tr>
<tr>
<td>Severe megacystis</td>
<td>5.16 (1.4 -18.6)</td>
<td>49%</td>
<td>84%</td>
</tr>
</tbody>
</table>

OR: odds ratio; CI confidence interval; GA gestational age
Table 3. Proposal of a clinical score for antenatal diagnosis of LUTO

<table>
<thead>
<tr>
<th></th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe megacystis (volume &gt; 35 cm³ or ascites)</td>
<td>4</td>
</tr>
<tr>
<td>Bilateral ureteral diameters</td>
<td>1.3 for each mm of ureteral size</td>
</tr>
<tr>
<td>Oligo or anhydramnios</td>
<td>4</td>
</tr>
<tr>
<td>Fetal sex (male)</td>
<td>4</td>
</tr>
<tr>
<td>Referral before the 28th week</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score ≥ 9.5</th>
<th>Risk of LUTO</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96%</td>
<td>78% (70-85)</td>
<td>79% (54-94)</td>
<td>96</td>
<td>36</td>
</tr>
</tbody>
</table>
Figure Legends

**Figure 1.** Study population

**Figure 2.** Receiver-operating characteristics curves for antenatal prediction of LUTO based on LUTO clinical score ( ), or based on the classic LUTO triad ( ).
Figure 1. Study population

Fetal Megacystis
n = 312

- Incomplete data on final diagnosis n = 98
- Incomplete data on prenatal US scans n = 68
- Not suspected for LUTO n = 3

Cases included
n = 143

- LUTO
  n = 124
- non-obstructive megacystis
  n = 19

This article is protected by copyright. All rights reserved.
Figure 2. Receiver-operating characteristics curves for antenatal prediction of LUTO based on LUTO clinical score ( ——), or based on the classic LUTO triad (----).

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>CI (95%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic LUTO triad</td>
<td>0.63</td>
<td>0.49 – 0.77</td>
<td>0.07</td>
</tr>
<tr>
<td>LUTO clinical score</td>
<td>0.84</td>
<td>0.75 – 0.93</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>