Direct Vescicoureteral Scintigraphy: Quantifying Early Outcome Predictors in Children with Primary Reflux

P. David Mozley, Sydney Heyman, John W. Duckett, Paul Turek, Howard Snyder, Greg Maislin and Harold Basheim

Division of Nuclear Medicine and Department of Urology, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, Pennsylvania

This study quantifies some of the outcome predictors in a group of children with primary vescicoureteral reflux who were initially managed medically. Methods: We studied 133 patients with primary reflex for 7.1 ± 2.2 yr. Direct vescicoureteral scintigraphy (DVS) was used to prospectively measure the absolute bladder volume at which reflex began and the maximum volume of urine refluxed into the ureters during the filling and voiding phases of their first two DVS studies. Findings were related to outcome as defined by spontaneous resolution or the eventual need for reconstructive surgery. Results: Medical management eventually failed in 35% of this sample. Patients who did not begin to reflux until their bladders had been filled to more than 60% total bladder capacity had a substantially smaller risk of surgery than those who began to reflux at smaller bladder volumes. Patients who refluxed a volume of urine back into their ureters that was less than about 2% of their total bladder capacity had a substantially smaller risk of surgery than those who refluxed more than 2%. The difference between groups was significant for both DVS variables (p < 0.001). Conclusion: Quantitative DVS contributes to the assessment of prognosis in children with vescicoureteral reflux who are managed medically.

Key Words: vescicoureteric reflux; prognosis; outcome


Vescicoureteral (VU) reflux is defined as the retrograde flow of urine from the bladder back into the ureters and renal collecting system (1). VU reflex remits spontaneously with growth in most children (2–4). In others, however, progressive renal scarring occurs (5–8) making reflux nephropathy one of the most common causes of chronic renal failure in children (9,10). VU reflux and the progression of its sequela can be stopped in most children with reconstructive surgery (11,12). However surgery carries several potential risks (13) and conservative medical management is satisfactory in most cases (14).

Several diagnostic procedures have been developed to identify children with reflux who may need to be triaged directly to surgery. Reflux does not appear to resolve spontaneously in many patients with severe, dilating forms of the disorder which correspond to a grade of 4 or 5 out of 5 on the radiographic voiding cystoureterogram (VCUG). On the other hand, reflux does resolve spontaneously without sequela in most (15), but definitely not all (16), cases of mild, nondilating reflux which correspond to a grade of 1 or 2 on the VCUG (17–20). Surgeons will usually operate on children with evidence of severe renal scarring on an intravenous pyelogram or a dimercapto-succinic acid (DMSA) scintigram (21–23). However, the need for surgery is still unclear after all of these procedures have been performed in some patients. These children tend to have low to intermediate grades of reflux, and are generally healthy.

The standard of care for following medically managed children with reflux who do not have a clear indication for surgery at the time of presentation has come to include periodic radionuclide cystography (24). The direct method of performing vescicoureteral scintigraphy (DVS) is a highly sensitive test for reflux (25–28) that can be performed with common radiopharmaceuticals (29) and essentially no biological risk to the patient (30,31). However, the DVS is frequently viewed as a yes-or-no test for continuing reflux that cannot be used to characterize the disorder any further (32). Absolute quantification is definitely feasible (33,34), but relatively labor intensive, and the pathophysiological significance of the many exact volume measurements that can be made has not yet been well established.

This study attempts to determine whether VU reflux could be characterized with quantitative scintigraphy. The inclusion criteria were designed to focus the investigation on a subgroup of patients with primary reflux who tended to be followed throughout childhood with multiple diagnostic procedures. The analysis was designed to determine the extent to which the reflux volume variables could predict the likelihood of conservative medical management succeeding or failing independently of the many other factors that influence decisions to abandon medical management and proceed to surgery.
METHODS

Patients

Patients were selected from hospital case records by identifying children who had their first direct voiding scintigram (DVS) between 1981 and 1984. Children with three or more reflux studies between 1981 and 1991 were included. Patients operated on within 2 yr of their first study were not included. Patients who were found to have another clinically significant disease process or congenital anomaly were also excluded. No other characteristics of the sample were known at the time of selection, including the results of the DVS.

The DVS was used to study 487 patients between 1981 and 1984. The selection process left 155 patients (34.2%) with a minimum of three DVS studies between 1981 and 1991. Measurements of reflux volume could not be recovered in 19 cases; records were missing on 12 patients. Technical problems prevented the data from being quantified in the others. Outcome could not be ascertained from either the hospital or the clinic records in 13 other children. The final sample consisted of 133 patients who were managed medically for at least 3 yr before being operated on or discharged from the clinic.

Outcome

Patients were followed by two relatively senior pediatric urologists practicing in a single setting. Resolution was defined by two consecutively negative DVS studies with a prolonged, symptom-free interval in between. The primary indications for abandoning medical management and proceeding to surgery before the onset of puberty included the development of pyelonephritis or renal scarring. Clinical assessments of general health, growth and development were documented and contributed to the process, but the weights that these factors were given could not be quantified. The surgeons reported that several psychosocial variables influenced their clinical decisions, including parental attitudes towards surgery and compliance with antibiotic prophylaxis (35). Several other factors that could not be quantified may have influenced the decisions to abandon medical management. However, quantifying the reflux volumes never entered into the process of making management decisions during this era. The DVS was used exclusively as a yes-or-no test for resolution. Hence, the reflux volumes did not influence the selection of the sample at all or have any impact on the final outcomes.

Radiological Studies

The analysis was confined to the initial grade of reflux on the first radiological voiding cystoureterogram (VCUG). The VCUG was performed by instilling radiographic contrast directly into the bladder under a fluoroscope. Reflux was graded according to the criteria proposed by the International Reflux Study Committee (36–38). Reflux confined to the ureter was classified as grade 1. Grade 2 designated reflux into the renal pelvis without dilatation of the ureter. Grades 3, 4 and 5 were used to indicate the degree of ureteral-pelviceal dilatation and tortuosity. The results were documented prospectively at the time of study. However, the findings were not known to any of the investigators who selected the sample.

The degree of renal scarring was assessed at the time of presentation in each patient. Measurements were made with a conventional intravenous pyelogram (IVP). The grading system was based on guidelines elaborated in a written protocol received from the International Reflux Study Committee in 1980. Severity was rated with a score of 0 if there was no radiographic evidence of scarring, and scores of 1, 2 or 3 were used to designate mild, moderate and severe renal scarring.

Direct Vescoureteral Scintigrams

The scintigraphic procedure began by preparing 500 ml of a normal saline solution containing 99mTc-pertechnetate. The exact dose of pertechnetate was measured in a dose calibrater (CRC-30, Capintec Inc., Montvale, NJ). The dose, which varied for each study, was based on an estimate or prediction of the maximum bladder capacity for each particular patient at the time of study. The estimate of total bladder capacity was calculated from the body surface area by assuming that an average human with a body surface area of 1.73 m² would have a maximum bladder volume of 500 ml (39). Up to 110 MBq (3 mCi) of pertechnetate were then mixed in 500 ml of normal saline so that the total dose delivered to each patient would equal 37 MBq (1 mCi) if the amount of saline infused equaled the estimated total bladder capacity. When the actual bladder capacity was higher than the estimated capacity, the infusion of pertechnetate was switched to an infusion of normal saline without radioactivity. The volume of nonradioactive saline was measured and an appropriate adjustment was made in the calculations of concentration.

The procedure was begun by catheterizing the urinary bladder with a 6 or 8 French feeding tube that did not have an indwelling balloon. Patients were never sedated. The catheter was taped in place after the residual urine was drained. The external end of the catheter was coupled to the reservoir of pertechnetate with plastic tubing. The reservoir was hung about 90 cm above the bladder. The solution containing the tracer was instilled at a constant rate designed to fill the bladder to its estimated capacity in 10–15 min. The rate of flow was tempered with an inclined roller clamp that cramped the plastic tubing.

Planar images were acquired in the posterior projection on clinical gamma cameras linked to a digital computer. Each image was acquired for 5 sec per frame during the filling phase of the study, and for 2 sec per frame during the voiding phase. Regions of interest (ROIs) were placed around the bladder, each ureter and the background. Background-corrected count rates were measured in each ROI on every frame. The frame with the maximum number of counts was determined independently for each ROI and used to calculate total bladder capacity and maximum volume of urine refluxed back into the collecting system during each phase of the study. The number of counts in the ROI for the bladder on the frame at which activity was first detected in a ureteral ROI was used to calculate the bladder volume at which reflux began on each side during each phase of the study.

The final concentration of activity in the system at the end of filling was calculated by dividing the total number of counts in the bladder and the ureters at the end of the filling phase by the volume of the solution that was infused. This measure of concentration was corrected for the amount of urine that was not aspirated out of the bladder before the study began and the volume of urine that was produced by the kidneys and let down into the collecting system during the study. Dilution from these sources was accounted for by comparing the count rates per unit volume in the infusion reservoir to the count rates per unit volume in an aliquot of the voided fluid.

Statistical Analysis

The results that were obtained from the first two DVS studies were analyzed independently. Subjects who never refluxed on either one of their first two DVS studies were already effectively excluded by the selection process. However, about a third of the
children in the sample did not reflux at all on one of their first two DVS studies. The subjects who did not reflux during the particular study being analyzed were segregated out of the sample as a separate group. The median value of each urine volume variable was determined for the remaining subjects who did reflux during the study. The reflux variables included the maximum volume of urine refluxed back into the collecting system, the bladder volume at which reflux first occurred, and the experimentally measured total bladder capacity.

The findings in each ureter were not analyzed independently, as if the two ureters were separate units that could have come from two different patients. The most extreme volume variables in each patient were used regardless of the side they occurred on. To prevent any of the reflux variables from biasing the selection of the other reflux variables, the maximum volume of urine refluxed back into the collecting system on each phase of the DVS was determined and used in the analysis, even if the maximums occurred on different sides. The bladder volume at which reflux first began on either side was used in the analysis, even if this did not correspond to the side with the maximum amount of urine refluxed back into the collecting system.

The value of each maximum or minimum volume variable was then used to split the patients who did reflux during the DVS into two more groups for each variable, depending on whether the value for a subject was more or less than the median value for the whole sample of patients who refluxed during that particular study. A univariate analysis of the correlation with outcome was performed for each demographic, clinical, radiographic and scintigraphic variable. The significance of the differences between groups in the failure rates of medical management were assigned using chi-square tests including tests for trends in proportions. The analyses of the scintigraphic data were repeated after normalizing the DVS volume variables for the maximum bladder capacity of each patient at the time of study.

A multivariable logistic regression was used to derive an equation which estimated the probability that surgery would eventually be required. It was based on the normalized volume variables from both the filling and the voiding phases, and the radiographic evidence of scarring. The volume variables from the first two studies were analyzed independently.

RESULTS

More than 80% (107 of 133) of the patients in this population were female. Their average age at the time of their first study was 3.98 ± 2.82 yr (median: 4.02 yr; range: 0.1–9.8 yr). They were followed for an average of 7.1 ± 2.2 yr (range 2.7–10 yr). The number of DVS studies before discharge or surgery ranged from an a priori set minimum of three to as many as 10. The average age at the time of surgery was 8.3 ± 3.1 yr (median 9.0 yr; range: 2.2–13.0 yr).

Reconstructive surgery was eventually performed on 47 of the 133 children (35%). Multiple statistical analyses showed that outcome was not related to the age of presentation, sex or the number of documented urinary tract infections during the entire follow-up period. There were no relationships between outcome and serum creatinine levels. There were no relationships between any of these variables and any of the reflux volumes, regardless of whether the volumes were normalized for the total bladder capacity.

The average grade of reflux on the initial radiographic VCUG was 2.2 in children who refluxed on the left and 2.4 in children who refluxed on the right. The grade of reflux on the first VCUG ranged from 0 (no reflux on the initial study) to 5. Despite this wide range of values, the grade of reflux on the initial VCUG was not significantly correlated with outcome in this highly selected population. A relationship could not be detected with either univariate or multivariate analyses. There were no relationships between the grade of reflux on the radiographic VCUG and the normalized reflux volumes in this particular sample.

Only 31% (41/133) of the patients in this sample had evidence of scarring. Of these, the grade was mild in about 71% (29/41) and moderate in the rest (12/41). None of the children in this population had evidence of severe scarring. The degree of scarring was not associated with age, gender or serum creatinine levels. However, the presence or absence of scarring was associated with outcome. Of the patients with any evidence of scarring at the time of presentation, 59% went on to surgery, but only 27% without scarring did (p < 0.0001). Differentiating between patients with mild renal scarring from patients with moderate renal scarring did not increase the predictive power of the measurements in any way.

There were no features of the initial DVS that correlated with outcome unless the volumes were normalized for the maximum bladder capacity. The patients who refluxed during the first DVS began to reflux during the filling phase when their bladders had been filled to a median of 60% of their total bladder capacity. Breaking the sample up into quartiles showed that 75% of the children who refluxed during the filling phase began to reflux by the time their bladders had been filled to 78% of the total bladder capacity. Twenty-five percent of the subsample who refluxed began to reflux by the time their bladders had been filled to 32% of the total bladder capacity.

Reflux began very early during the voiding phase in most children. Half the children who refluxed during the voiding phase had already begun to reflux before their bladders had contracted down to a volume that was 95% of their total bladder capacity. The maximum amount of urine refluxed from the bladder back into the collecting system had a median volume of 2.1% ± 0.3% of the total bladder capacity during the filling phase and a median volume of 1.9% during the voiding phases of the first study.

All patients refluxed on at least one of their first two DVS studies. However, 30% (43/142) of the patients did not reflux during either the filling or voiding phases of their very first study. Patients who did not reflux during their initial DVS study had the smallest risk of surgery (range: 6%–12%). Patients who began to reflux after their bladders had been filled to more than 60% of their bladder capacity had a smaller risk of surgery than those who began to reflux at lower bladder volumes. Patients who refluxed a volume of urine back into a ureter that was less than the median
volume for the entire sample had a substantially smaller risk of surgery than those who refluxed more than the median volume. However, this finding was only significant during the voiding phase of the first study (Chi-square for reflux >1.9% on voiding phase = 8.3, df = 2, p = 0.02; Chi-square for filling phase = 2.0, df = 2, p = 0.37). The differences in the risk of surgery as a function of both scintigraphic variables taken together was significant (p < 0.0001).

The analysis was repeated with the volume measurements that were acquired during the second DVS study. The maximum volume of urine refluxed back into the ureters of children who refluxed during the second study had a median of 1.9% of the total bladder capacity, beginning during the filling phase when their bladders had been expanded to a median of 56% of the total bladder capacity. On the voiding phase, most children who refluxed began to reflux very early, when their bladders had contracted down to a median volume of only 95% of their total bladder capacity. The predictive power of both volume variables was significant on both the filling and voiding phases of the second study.

A multivariable logistic regression was used to derive an equation which estimated the probability that surgery would ultimately be required to manage reflux. It was based on the presence of scarring at the time of presentation and the severity of reflux on the first follow-up exam. The Hosmer-Lemeshow test (40) led to accepting the null hypothesis that the model fit the data adequately (chi-square = 4.0, df = 8, p = 0.86). As measured by the Gamma statistic (41) the overall predictive ability of the model appeared good (g = 0.57). The estimated probability that medical management would fail ranged from 0.06 to 0.12 in patients without scarring who did not reflux on their first scintigram to 0.70 to 0.92 for patients with scarring who refluxed a volume of urine back into their ureters that was more than 1.7% or 2.1% of their bladder capacity on both the voiding and filling phases of the study. The results are summarized in Tables 1 and 2 for the first DVS and Tables 3 and 4 for the second DVS.

**DISCUSSION**

This study attempts to quantify some of the early predictors of outcome in a subgroup of children with primary VU reflux. The results indicate that the bladder volume at which reflux began and the volume of urine that was refluxed back into the collecting system on the first two DVS studies could have contributed to the assessment of prognosis in this group. The disorder was less likely to resolve spontaneously in children who began to reflux at a relatively lower bladder volume during the filling phase. Children who refluxed relatively more urine back into the collecting system were more likely to need reconstructive surgery than children who refluxed relatively less. Medical management was more likely to fail in children who refluxed during both phases of the DVS than in children who only refluxed during one phase. When combined with radiographic evidence of any scarring, the reflux volume variables could account for between 86%-92% of the variance in the model which described the probability of medical management eventually failing.
Reflux begins at filing phase of blender volume when DVS > @% of TBC and TBC < 60% of TBC. No renal scarring at time of presentation.

<table>
<thead>
<tr>
<th></th>
<th>No reflux on the filing phase of the first DVS</th>
<th>Reflux begins at bladder volume &gt;60% of TBC</th>
<th>Reflux begins at bladder volume &lt;60% of TBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reflux on the voiding phase of the first DVS</td>
<td>0.12</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &lt;95% of the TBC</td>
<td>0.19</td>
<td>0.22</td>
<td>0.34</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &gt;95% of the TBC</td>
<td>0.31</td>
<td>0.35</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Renal scarring found on initial exam.

<table>
<thead>
<tr>
<th></th>
<th>No reflux on the voiding phase of the first DVS</th>
<th>Reflux begins at bladder volume &gt;60% of TBC</th>
<th>Reflux begins at bladder volume &lt;60% of TBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reflux on the voiding phase of the first DVS</td>
<td>0.40</td>
<td>0.42</td>
<td>0.58</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &lt;95% of the TBC</td>
<td>0.49</td>
<td>0.52</td>
<td>0.67</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &gt;95% of the TBC</td>
<td>0.74</td>
<td>0.76</td>
<td>0.86</td>
</tr>
</tbody>
</table>

TBC = total bladder capacity.

The selection process was designed to focus the investigation on a subgroup of patients with primary reflux who were initially managed medically. Most of the children who were included were followed throughout much of their childhood with multiple diagnostic procedures. Children who were operated on within 2 yr of their first DVS were excluded. Most of the patients who quickly went to surgery probably presented with very high grades of reflux or evidence of severe renal scarring. This is consistent with the finding that none of the patients in the study population had evidence of severe renal scarring on their initial IVP. Eliminating children who may have had more severe forms of the disorder should have made it more difficult to find an association between the reflux volumes and outcome by

<table>
<thead>
<tr>
<th></th>
<th>No reflux on the filing phase of the second DVS</th>
<th>Reflux volume on the filling phase &lt;2% of the TBC</th>
<th>Reflux volume on the filling phase &gt;2% of the TBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reflux on the voiding phase of the second DVS</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>Reflux volume on the voiding phase is &lt;2% of TBC</td>
<td>0.21</td>
<td>0.35</td>
<td>0.49</td>
</tr>
<tr>
<td>Reflux volume on the voiding phase is &gt;2% of TBC</td>
<td>0.26</td>
<td>0.42</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Renal scarring found on initial exam.

<table>
<thead>
<tr>
<th></th>
<th>No reflux on the voiding phase of the second DVS</th>
<th>Reflux volume on the filling phase &lt;2% of the TBC</th>
<th>Reflux volume on the filling phase &gt;2% of the TBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reflux on the voiding phase of the second DVS</td>
<td>0.27</td>
<td>0.42</td>
<td>0.57</td>
</tr>
<tr>
<td>Reflux volume on the voiding phase is &lt;2% of TBC</td>
<td>0.53</td>
<td>0.69</td>
<td>0.80</td>
</tr>
<tr>
<td>Reflux volume on the voiding phase is &gt;2% of TBC</td>
<td>0.60</td>
<td>0.75</td>
<td>0.85</td>
</tr>
</tbody>
</table>

TBC = total bladder capacity.
compressing the range of severities in the study population.

The range of severities may also have been compressed by excluding children who were not followed with at least three DVS studies. Most of the patients in the subgroup probably had very low grades of reflux and no evidence of scarring. The radiographic VCUG has already been shown to be useful for predicting outcome in children with very mild and very severe reflux. This investigation attempted to study children who tend to require prolonged follow-up. Compressing the range of severities in the study population suggests that the capacity of the quantitative DVS to contribute to assessments of prognosis may be even higher in the whole population of children with reflux. The fact that the radiographic VCUG did not correlate with outcome in this particular population tends to support this suggestion. The grade of reflux did have a correlation with outcome in a more heterogeneous group of patients from this hospital who were evaluated during the same time by the same urologists (35). Likewise, the degree of renal scarring did not have as much predictive power in this sample as it has in others that used similar radiographic techniques to measure scarring. But, just as the predictive power of conventional radiographic procedures may have been limited by the narrowness of the inclusion criteria in this sample, the predictive power of the quantitative DVS may be even greater in the whole spectrum of patients with minimal to extremely severe reflux.

The results seem to provide some general support for an earlier study which indicated that children who refluxed at a higher, absolute bladder volume on follow-up had a better prognosis than children who refluxed at a lower, absolute bladder volume (33). The absolute bladder capacity tends to increase with age in children. It may be that the changes which were observed in that study simply tend to corroborate our findings that children who began to reflux at a relatively lower bladder volume did better than children who began to reflux at a relatively higher volume.

The predictive power of any diagnostic imaging procedure will probably always be limited by the many other factors that contribute to the decision to intervene surgically in these patients (42). Some of these factors are complex and difficult to control, such as assessments of general health, growth and development. Other potentially important variables may not be directly related to the pathophysiology of the disorder, such as parental attitudes towards surgery and compliance with antibiotic prophylaxis (35,42). Difficulties quantifying all the variables that influenced the decisions to abandon medical management should have made it harder to find an association between outcome and the reflux volume variables. This may partially explain why almost 1 out of 10 children with no evidence of renal scarring at the time of presentation who did not reflux on their initial DVS exam still eventually needed reconstructive surgery. Nevertheless, the findings in this investigation are strong enough to suggest that DVS may reduce some of the uncertainty surrounding the management of some patients with reflux who are initially managed medically.

**TABLE 4**
Estimates Based on Bladder Volume where Reflux Begins During One-Year Follow-up DVS

<table>
<thead>
<tr>
<th>No reflux on the filling phase of the second DVS</th>
<th>Reflux begins at bladder volume &gt;60% of TBC</th>
<th>Reflux begins at bladder volume &lt;60% of TBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reflux on the voiding phase of the second DVS</td>
<td>0.06</td>
<td>0.11</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &lt;95% of the TBC</td>
<td>0.11</td>
<td>0.18</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &gt;95% of the TBC</td>
<td>0.31</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Renal scarring found on initial exam

<table>
<thead>
<tr>
<th>No reflux on the voiding phase of the second DVS</th>
<th>0.25</th>
<th>0.38</th>
<th>0.63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflux begins at a bladder volume of &lt;95% of the TBC</td>
<td>0.37</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td>Reflux begins at a bladder volume of &gt;95% of the TBC</td>
<td>0.69</td>
<td>0.81</td>
<td>0.92</td>
</tr>
</tbody>
</table>

TBC = total bladder capacity.
REFERENCES


EDITORIAL

The Ongoing Challenge of Diagnosis and Treatment of Urinary Tract Infection, Vesicoureteral Reflux and Renal Damage in Children

In this issue of the Journal, Mozley et al. provide an opportunity to comment on several aspects surrounding the diagnosis and treatment of vesicoureteral reflux (VUR) and urinary tract infection (UTI) as well as the identification of renal damage.

The goals of the physician taking care of children with urinary tract infection are to avoid renal damage if possible, or, if damage has already occurred, to prevent it from progressing further (1-5). Renal scarring can lead to growth failure, hypertension and even chronic renal failure (6,7). When the diagnosis of UTI in children is made (clinical history, general appearance of the child, urine analysis, urine culture, etc.), several questions