A scoring system for predicting downgrading and resolution of high-grade infant vesicoureteral reflux

Sofia Sjöström1,2,3 | Aldina Pivodic4,5 | Kate Abrahamsson1,2 | Rune Sixt1,6 | Eira Stokland1,7 | Sverker Hansson1,8

Abstract

Aim: Aim of the study was to provide a scoring system for predicting downgrading and resolution of infantile high-grade vesicoureteral reflux (VUR).

Methods: Eighty-nine infants (65 boys) with high-grade VUR (grade 4-5) diagnosed at median age 2.5 months and followed to 39 months had repeated investigations of VUR grade, renal damage/function and bladder function. Recurrent urinary tract infections (UTIs) were registered. Risk variables collected at 1 year were analysed as independent factors for spontaneous resolution to grades ≤2 and 0, using univariable/multivariable logistic regression.

Results: A scoring system was built with a total of 14 points from four independent risk factors (sex, breakthrough UTI, type of renal damage and subnormal glomerular filtration rate). Children with persistent VUR (grade 3-5) had higher scores compared with the group with spontaneous resolution (grade 0-2) (mean 7.9 vs. 4.5, P < .0001). A score of ≥8 points indicated a low probability of VUR resolution (≤14%). The model was considered excellent based on area under the ROC curve (0.82) and showed satisfactory internal validity.

Conclusion: This model provides a practical tool in the management of infants born with high-grade reflux. High scores at one year of age indicate a high risk of persistent dilated reflux.

KEYWORDS

high-grade, infants, resolution and downgrading, scoring system, vesicoureteral reflux
Vesicoureteral reflux (VUR) is a common finding in a paediatric population and occurs in about 1%-2%, with prevalence inversely proportional to age, as many resolve spontaneously. The natural course of VUR depends on grade, age, sex, concurrent bladder dysfunction and many other factors that have been shown to be negative predictors of the spontaneous resolution of VUR.1-3 From these known risk factors, a number of studies have suggested models for predicting VUR resolution.4-7 Most of these studies include children with all grades of VUR and with high-grade VUR in a minority. High-grade VUR showed poor prognosis of resolution in these mixed populations.

Since the overall aim of reducing the risk of recurrent febrile urinary tract infection (UTI) and new renal damage is driving the clinical treatment of VUR patients, risk factors for recurrent febrile infections have also been studied. Algorithms focusing on the risk of UTI for use in the clinical setting have been created. Not surprisingly, the risk factors that were found were much the same as for the non-resolution of VUR: high-grade VUR, age, female sex, reflux at low bladder volume, bladder bowel dysfunction and UTI history.8,9

Children younger than two years have a higher resolution rate than older children, if all VUR grades are included.5,10 An index for predicting resolution in this group has been suggested, including sex, ureteral anomalies, high-grade VUR and the timing of VUR occurrence as variables.6,7 These studies also included children with high-grade VUR, but it appears that they were censored at the end of follow-up, as they did not resolve or were surgically treated, that is time to event was missing. However, previous studies of infants with high-grade VUR have shown downgrading and spontaneous resolution of reflux, even in this group, but to a lesser extent than for lower grades.11 The downgrading or resolution has mainly been seen in grade 4 and unilateral grade 5 but seldom in bilateral grade 5 in a one-year perspective.12

Infants with high-grade VUR have higher morbidity than infants with low-grade VUR, due to the complexity of the disease. In these patients, VUR is often combined with massive dilation of the upper urinary tract, generalised renal damage, concomitant bladder dysfunction and a high risk of UTI. Despite this, there is a subgroup with early resolution, seen already during the first years of life,11 and therefore predicting the probability of VUR resolution in the individual child would be beneficial before deciding on treatments.

In this study, the objective was to develop a pragmatic tool for predicting VUR resolution and downgrading (to grade ≤2) at a given time, in high-grade reflux presenting during infancy. The study aimed to construct a scoring system from variables shown to have significant independent effects on VUR resolution in infants.3 From these variables, we aimed to provide a sum score with cut-off values, selecting patients with a high, intermediate and low probability of spontaneous resolution of reflux and thus to select patients in need of more active VUR treatment.

### Key notes
- Infants with high-grade vesicoureteral reflux (VUR) have low spontaneous resolution rate, even if a small proportion without risk factors has shown resolution (eg without renal damage, recurrent urinary-tract-infections and bladder dysfunction).
- A scoring system was developed, with a good accuracy of predicting VUR resolution.
- This model can be a practical tool for making decisions on further treatment, surveillance, prophylaxis or surgical intervention, in children with high-grade VUR at age one-year.

### 2 | MATERIALS AND METHODS

#### 2.1 | Patient material

Eighty-nine infants (65 boys) with high-grade VUR (grade 4-5) diagnosed during infancy (median age 2.5 months) provided the data for this study. The data were taken from a previously reported prospective observational study of a cohort of infants diagnosed with high-grade VUR during their first year of life.11 Predictive factors for the spontaneous resolution of VUR or risk factors for remaining diluted VUR in the same cohort have also been presented.3 Compared with the original study (n = 115), we chose to include only VUR grade 4 and 5 for calculations of the scoring system in the present study. The excluded patients from the original study were 18 children with VUR grade 3 and eight with resolution/downgrading to grade ≤2 already during the first year of life. Characteristics of the new study population are shown in Table 1.

#### 2.2 | Study design

The patients were followed for a median of 39 months according to a study protocol with repeated examinations. The first investigation was performed after the diagnosis of VUR and then repeated yearly, or until VUR resolution (grade 0) or downgrading to VUR grade ≤2 or until the performance of an anti-reflux procedure. In the study, two endpoints were evaluated; first the number of patients with VUR resolution/downgrading to grade ≤2 (grades 0,1,2) and second, number of patients with complete VUR resolution (grade 0). To determine VUR grade (videocystometry, VCM), bladder function (VCM and four-hour voiding observation), renal abnormalities (DMSA and MAG3 scintigraphy) and renal function (glomerular filtration rate (GFR) by $^{51}$Cr-EDTA clearance or estimated from Schwartz formula) were performed in each patient. All the children were kept on antibacterial prophylaxis and recurrent infections (defined as the growth of at least 100,000 colony forming units per ml urine and fever of at least 38.5°C or higher) were registered. Investigation
| Variable                        | Total (n = 89) | Grade IV (n = 47) | Grade V (n = 42) |
|--------------------------------|----------------|-------------------|------------------|
| **Age at first VCUG (months)** | 3.5 (3.5)      | 4.0 (3.3)         | 2.9 (3.5)        |
|                                | 2.5 (0.03-11.7)| 2.8 (0.1-11.7)    | 1.0 (0.03-10.9)  |
| **Age at one-year follow-up (months)** | 15.1 (5.8) | 15.3 (5.2) | 14.8 (6.3) |
|                                | 14.9 (1.9-23.7)| 15.8 (4.0-23.7)  | 14.6 (1.9-23.5)  |
| **Age at last follow-up (months)** | 37.3 (15.0) | 36.5 (15.1) | 38.2 (15.1) |
|                                | 39.0 (11.7-68.6)| 37.0 (12.1-68.6)| 39.9 (11.7-65.7)|
| **Gender**                     |                |                   |                  |
| Male                           | 65 (73%)       | 29 (62%)          | 36 (86%)         |
| Female                         | 24 (27%)       | 18 (38%)          | 6 (14%)          |
| **Presentation of VUR**        |                |                   |                  |
| Prenatal ultrasound           | 24 (27%)       | 8 (17%)           | 16 (39%)         |
| UTI                            | 64 (73%)       | 39 (83%)          | 25 (61%)         |
| **Uni/bilateral VUR**         |                |                   |                  |
| Unilateral                     | 25 (28%)       | 17 (36%)          | 8 (19%)          |
| Bilateral                      | 64 (72%)       | 30 (64%)          | 34 (81%)         |
| **Breakthrough UTI**          |                |                   |                  |
| No                             | 52 (58%)       | 31 (66%)          | 21 (50%)         |
| Yes                            | 37 (42%)       | 16 (34%)          | 21 (50%)         |
| **Number of breakthrough UTIs**| 0.8 (1.2)     | 0.5 (0.9)         | 1.1 (1.5)        |
|                                | 0 (1-5)        | 0 (0-3)           |                  |
| **Renal abnormality**         |                |                   |                  |
| No                             | 8 (9%)         | 3 (6%)            | 5 (12%)          |
| Yes                            | 81 (91%)       | 44 (95%)          | 37 (88%)         |
| **Renal abnormality**         |                |                   |                  |
| None                           | 8 (9%)         | 3 (6%)            | 5 (12%)          |
| Focal                          | 22 (25%)       | 13 (28%)          | 9 (21%)          |
| Generalised                    | 59 (66%)       | 31 (66%)          | 28 (67%)         |
| **Uni/bilateral renal abnormality** | 8 (9%) | 3 (6%) | 5 (11%) |
| None                           | 62 (70%)       | 37 (79%)          | 25 (60%)         |
| Unilateral                     | 19 (21%)       | 7 (15%)           | 12 (29%)         |
| **Bladder dysfunction**       |                |                   |                  |
| No                             | 24 (28%)       | 14 (32%)          | 10 (24%)         |
| Uncertain                      | 37 (44%)       | 21 (48%)          | 16 (39%)         |
| Yes                            | 24 (28%)       | 9 (20%)           | 15 (37%)         |
| **Bladder capacity (% of expected for age)** | 190 (104) | 186 (110) | 194 (98) |
|                                | 191 (29.467) | 176 (29.467)      | 198 (37.404)     |
| **Bladder capacity ≥150%**    |                |                   |                  |
| No                             | 36 (41%)       | 20 (44%)          | 16 (38%)         |
| Yes                            | 51 (58%)       | 25 (56%)          | 26 (62%)         |
| **Residual urine (% of bladder capacity)** | 26 (17) | 22 (15) | 30 (18) |
|                                | 23 (0-90)      | 19 (0-67)         | 27 (0-90)        |
| **GFR categorised**           |                |                   |                  |
| Normal                         | 48 (55%)       | 31 (66%)          | 17 (41%)         |
| Subnormal                      | 40 (45%)       | 16 (34%)          | 24 (59%)         |

Note: For categorical variables, n (%) is presented.
For continuous variables, the mean (SD)/median (min; max)/n = is presented.
results for prediction analyses were taken from a cross-section at approximately one year of age in each of the included patients.

2.3 | Investigation methods

VCM includes simultaneous voiding cystourethrography (VCUG) and cystometry and has previously been described. The grading of VUR was taken from the VCUG part of the VCM, was performed according to international standards and, in bilateral cases, the patients were categorised by the side with the highest reflux. Bladder capacity was estimated from the highest sum of voided volume and residual urine in the cystometric part. Bladder capacity of ≥150% of expected for age was defined as large and expected for age was calculated [30*(age [years] + 1)] ml.

FVO is an assessment of the free voiding pattern. Residual urine was estimated as the mean of the residuals registered during the observation, measured with ultrasound. Elevated residual means ≥10 ml, according to FVO investigations in healthy, non-toilet-trained children, but, in the present study, ≥25% of bladder capacity was used due to the large variation in bladder capacity. Lower urinary tract dysfunction (LUTD) was defined as both large bladder capacity (≥150% of expected) and increased residual urine (≥25% of bladder capacity).

DMSA (dimercaptosuccinic acid) and MAG3 (mercaptoacetyltriglycerine) scintigraphy were performed in accordance with European guidelines. Focal renal damage was defined as area(s) with reduced tracer uptake or indentation of the renal outline, while generalised damage was defined as a small kidney with reduced tracer uptake or diffuse parenchymal abnormality and with a split function of less than 45%.

The GFR (\(51^{\text{Cr}}\text{-EDTA clearance or estimated clearance}\)) reference value was taken as a mean 110 ml/minute per 1.73 m\(^2\) body surface area for children after two years of age and according to the equation of Winberg below two years of age (Appendix S1). Subnormal GFR was arbitrarily defined as <80% of normal (~2SD; ie <88 ml/minute per 1.73 m\(^2\) after 2 years of age). The limit chosen for decreased GFR in this study was similar to definition of chronic kidney disease (CKD) stage 2 or higher (GFR <89 ml/min per 1.73 m\(^2\)). In patients, where \(51^{\text{Cr}}\text{-EDTA clearance was missing estimated clearance was calculated according to an equation like the modified Schwartz formula using values of body length and plasma creatinine.}}

2.4 | Statistical methods

For descriptive purposes, the mean, standard deviation (SD), median, minimum and maximum, or interquartile range (IQR) where applicable were presented for continuous variables, while the number and percentage were presented for categorical variables. For tests between groups, the Mann-Whitney U test was used for continuous variables, Mantel-Haenszel chi-square test for ordered categorical and Fisher’s exact test for dichotomous variables.

The prediction of VUR ≤ 2 and VUR = 0 was performed using logistic regression. All potential predictors were first tested separately in univariable models. Stepwise logistic regression was then performed to select a set of independent statistically significant predictors in a multivariable model, allowing entrance of variables having \(P < .10\) in the univariable regression. For descriptive purposes, estimates, standard errors (SEs), odds ratios (ORs) with 95% confidence intervals (CIs) and \(P\)-values were presented from the model, along with the area under the ROC curve (AUC) describing the model fit. The AUC between 0.70 and <0.80 is interpreted as acceptable, 0.80 and <0.90 as excellent and ≥0.90 as outstanding. The scoring value for each independent statistically significant predictor was defined based on the estimates obtained from the multivariable analysis, ie the weights of the predictor adjusted for each other. The probability of having VUR ≤ 2 based on the total score was calculated using the following algorithm:

\[
\text{Probability} = \frac{\exp(lc)}{1 + \exp(lc)},
\]

where

\[lc = \text{linear combination} = \text{estimate}_{\text{intercept}} + (\text{total score}) \times \text{estimate}_{\text{total score}}\]

The internal validation of the model was performed by using non-parametric bootstrapping technique, reflecting the process of sampling from the underlying population. The bootstrap samples were drawn with replacement from the original population, sampling 5,000 studies of the same size as the original sample (\(n = 89\)). The stepwise logistic regression analyses were then performed on each bootstrapped study including 89 patients each and the best model was chosen. The frequency of these models was counted and the mean, median, 25th percentile and 75th percentile of the area under the ROC curve were summarised for all 5,000 studies.

Sensitivity, specificity and positive predicted values (PPV) were presented for selected cut-offs.

One patient (1.1%) had missing data in the multivariable model for prediction of VUR ≤ 2 that was not imputed.

All the tests were two-tailed and conducted at the 5% significance level. All the analyses were performed in SAS® version 9.4 (SAS Institute Inc, Cary, NC, USA).

Ethical approval for the original study was granted by the committee of ethics for medical research at the University of Gothenburg (No Ö 140-03).

3 | RESULTS

3.1 | Study population

The results of investigations, from a cross-section at approximately one year of age (median 14.9 months [interquartile range (IQR) 11.8-19.8]), were chosen in each of the 89 included patients. Description of the cohort at one year is shown in Table 1, with almost equal distribution between grades 4 and 5 and with bilateral VUR in 72%. Renal abnormality was seen in 91%, with generalised damage in two
thirds of the patients. GFR was found subnormal (<80% of expected) in 46%, measured by $^{51}$Cr-EDTA clearance in 80 (90%), estimated from Schwartz formula in 8(9%) and missing in one patient. LUTD was seen in about one third and breakthrough UTI during the first year of life in 42%. Some differences between grade 4 and grade 5 were observed, where those with grade 5 VUR were younger at diagnosis, more of male sex, overrepresented regarding prenatal diagnosis, subnormal GFR and increased post-void residual urine.

Median time from VCUG close to 1 year to resolution/downgrading of VUR to grade 2 was 29.8 (IQR 16.1-44.8) months where 27 children (30%) had VUR resolution/downgrading to grade ≤ 2 and of these, 19 (21%) had complete resolution (grade 0).

### 3.2 Prediction of VUR resolution/downgrading to grade ≤ 2

Applying univariable logistic regression, the variables included in Table 2 were assessed for the prediction of VUR ≤ 2 at the latest follow-up for each child in order to identify parameters appropriate for inclusion in the multivariable models.

The following variables were identified as predictors of VUR ≤ 2: male sex, absence of breakthrough UTI, absence of bladder dysfunction, low bladder capacity (≤150%), low post-void residual urine and normal GFR (≥80% of expected for age). Additionally, unilateral VUR and absence of, or focal renal abnormality showed borderline predictive ability, fulfilling the requirements for entering the stepwise logistic regression (Table 2). From these items, the best multivariable model selected sex, breakthrough UTI, renal damage categorised and GFR categorised (Table 2). In this model, the area under the ROC curve (AUC) was 0.82 (95% CI 0.72-0.92), that is indicating a strong model (Figure 1).

### 3.3 Score for VUR resolution/downgrading to grade ≤ 2

Based on the estimates obtained from the selected multivariable logistic model the scoring was developed for each of the four variables resulting in a sum score ranging between 0 and 14 (Table 3). Figure 2 presents the probability curve with confidence intervals for the constructed sum score and describes the algorithm. The AUC for the model including only the score was 0.82 (95% CI 0.72-0.91) (Figure 1).

The estimated probability for VUR resolution to grade ≤ 2 ranged between 91.2% (95% CI 69.5%-97.9%) for score 0 to 0.7% (95% CI 0.1%-6.2%) for score 14 (Figure 2).

The sum of scores in those with spontaneous resolution/downgrading to grade ≤ 2 was a mean of 4.5 (SD 2.67) compared with 7.9 (SD 2.73) in the group with persistent VUR (P = .0001), (Table 4). A score of eight points or more resulted in low probability of VUR resolution (≤14%) for each score level, mean 6.7% [95% CI 0.8%-22.1%], (Table 5). Using this cut-off, the sensitivity for correctly identify...
TABLE 2 Variables identified in univariate logistic regression analyses as predictors to VUR ≤ 2, and thus appropriate for inclusion in the multivariable models. The best multivariable model selected the variables shown in the table.

| Variable                              | Value      | n (%) of event | Univariable OR (95% CI) | P-value | Area under ROC-Curve (95% CI) | Multivariable OR (95% CI) | P-value |
|---------------------------------------|------------|----------------|-------------------------|---------|-----------------------------|--------------------------|---------|
| Sex                                   | Boy (1)    | 24 (37%)       | 0.24 (0.07-0.90)        | .035    | 0.61 (0.53-0.70)            | 0.13 (0.02-0.68)         | .016    |
|                                       | Girl (2)   | 3 (13%)        |                         |         |                             |                          |         |
| Presentation of VUR                   | Prenatal US | 5 (21%)        | 1.86 (0.61-5.66)        | .28     | 0.56 (0.46-0.65)            |                          |         |
|                                       | UTI        | 21 (33%)       |                         |         |                             |                          |         |
| Grade of VUR at inclusion (IV-V)      | IV         | 17 (36%)       | 0.55 (0.22-1.39)        | .21     | 0.57 (0.46-0.68)            |                          |         |
|                                       | V          | 10 (24%)       |                         |         |                             |                          |         |
| Uni/Bilateral reflux                  | Unilateral | 11 (44%)       | 0.42 (0.16-1.12)        | .084    | 0.59 (0.48-0.70)            |                          |         |
|                                       | Bilateral  | 16 (25%)       |                         |         |                             |                          |         |
| Breakthrough UTI                      | No (0)     | 22 (42%)       | 0.21 (0.07-0.63)        | .0055   | 0.67 (0.57-0.76)            | 0.21 (0.06-0.71)         | .012    |
|                                       | Yes (1)    | 5 (14%)        |                         |         |                             |                          |         |
| No of breakthrough UTIs               | 0-<1       | 22 (42%)       |                         |         |                             |                          |         |
|                                       | 1-<2       | 2 (10%)        |                         |         |                             |                          |         |
|                                       | 2-<3       | 1 (17%)        |                         |         |                             |                          |         |
|                                       | 3-<4       | 2 (33%)        |                         |         |                             |                          |         |
| Renal abnormality                     | No         | 5 (63%)        | 0.22 (0.05-1.02)        | .052    | 0.57 (0.49-0.65)            |                          |         |
|                                       | Yes        | 22 (27%)       |                         |         |                             |                          |         |
| Renal abnormality                     | None (0)   | 5 (63%)        |                         |         |                             |                          |         |
|                                       | Focal (1)  | 7 (32%)        |                         |         |                             |                          |         |
|                                       | Generalised (2) | 15 (25%) | 0.52 (0.27-1.02) | .059 | 0.59 (0.48-0.71) | 0.32 (0.13-0.82) | .017 |
| Uni/Bilateral renal abnormality       | None       | 5 (63%)        |                         |         |                             |                          |         |
|                                       | Unilateral | 18 (29%)       |                         |         |                             |                          |         |
|                                       | Bilateral  | 4 (21%)        | 0.43 (0.17-1.07)        | .069    | 0.60 (0.49-0.70)            |                          |         |
| Bladder dysfunction                   | No         | 12 (50%)       |                         |         |                             |                          |         |
|                                       | Uncertain  | 10 (27%)       |                         |         |                             |                          |         |
|                                       | Yes        | 4 (17%)        | 0.44 (0.22-0.85)        | .015    | 0.66 (0.54-0.78)            |                          |         |
| Bladder capacity (% of expected/10)  | 2.9-<10.4  | 9 (39%)        |                         |         |                             |                          |         |
|                                       | 10.4-<20   | 8 (35%)        |                         |         |                             |                          |         |
|                                       | 20-<30.8   | 6 (23%)        |                         |         |                             |                          |         |
|                                       | 30.8-46.7  | 3 (20%)        | 0.95 (0.91-1.00)        | .063    | 0.62 (0.49-0.75)            |                          |         |
| Bladder capacity (%)                  | <=150      | 15 (42%)       | 0.39 (0.15-0.99)        | .047    | 0.62 (0.50-0.73)            |                          |         |
|                                       | >150       | 11 (22%)       |                         |         |                             |                          |         |
hand, this is not always necessary. The inclusion of renal damage in our scoring system is therefore justified, especially since the absence of renal damage in high-grade VUR children also previously has been shown to predict the resolution of VUR.\textsuperscript{1-3} Furthermore, in a study of a cohort of 161 children of different ages and all grades of VUR (only 12% high-grade VUR included), a computational model showed that, when adding renal scan data, the accuracy of the model improved.\textsuperscript{4,22}

Regarding estimates of renal function as compared to measured GFR, the latter investigations may not be routine at all institutions. If so, estimated GFR calculated from the formula published by Schwartz (see methods) from information on serum creatinine and body length/height can be used.\textsuperscript{13} This estimation is not as exact as assessments with for example 51Cr EDTA, but it is most often sufficient to identify patients with severely impaired renal function. For the proposed score, only information on normal and subnormal (<80% of normal) function is required. Another obstacle is that GFR changes during the first two years and reach adult levels (110 ml/min/1.73 m\textsuperscript{2}) at approximately two years of age. Winberg's logarithm regarding GFR in healthy infants is therefore recommended for the calculation of age-dependent reference values, enclosed in a Table S1 to make calculations easier.\textsuperscript{19}

It is well known that VUR in general is more common in girls but high-grade infant reflux has a preponderance among males. Sex was also an important factor in predicting reflux resolution in this study, with poorer prognosis in girls with high-grade VUR in this age group. Girls have been shown to have slower rates of resolution, especially in cases of bilateral VUR, both in studies including all ages\textsuperscript{5} and in studies specifically assessing infants with high-grade VUR.\textsuperscript{11} It can be hypothesised that the slower resolution in girls in this age group is associated with the higher risk of recurrent UTIs. This was supported by the results in the Swedish Reflux Trial.\textsuperscript{23} In this study, a

| Variable | Value | n (%) of event | OR (95% CI) VUR = 0,1,2 | P-value | Area under ROC-Curve (95% CI) | OR (95% CI) VUR = 0,1,2 | P-value |
|----------|-------|----------------|-------------------------|---------|-------------------------------|-------------------------|---------|
| Post-void Residual urine | 0-<1 | 8 (57%) | | | | | |
| | 1-<2 | 8 (33%) | | | | | |
| | 2-<3 | 6 (38%) | | | | | |
| | 3-<4 | 1 (7%) | | | | | |
| | 4-9 | 3 (18%) | 0.69 (0.49-0.97) | .031 | 0.67 (0.54-0.79) | | |
| GFR categorized | Normal (0) | 20 (42%) | 0.30 (0.11-0.81) | .017 | 0.64 (0.54-0.75) | 0.22 (0.07-0.71) | .011 |
| | Subnormal (1) | 7 (18%) | | | | | |

**TABLE 3** Scoring points for each predictor in forecasting the probability of resolution/downgrading to grade ≤2 of infantile high-grade VUR, estimated from data at age one year. Points obtained from estimates (weight of variable) in multivariable logistic regression analyses

Abbreviations: GFR, glomerular filtration rate; UTI, urinary tract infection.
higher risk of recurrent UTIs and new renal damage was seen in girls, but not in boys, in the surveillance group in children with grades 3-4 VUR at age one to three years. Interestingly, this was mainly seen in girls with non-resolution of reflux during follow-up.

Infants with high-grade VUR have a high risk of recurrent UTI despite continuous antibiotic prophylaxis. The frequency of breakthrough UTI in this study was as high as 42% but comparable to other cohorts with a high proportion of uncircumcised infant males. Several studies have reported breakthrough UTI in 46%-59% of children with high-grade VUR and with higher frequency in girls than boys, especially after the first year. Infants and children with high-grade VUR also have a higher risk of renal scarring after a febrile UTI. In view of the importance of minimising febrile recurrent UTIs in order to reduce the risk of progressive renal damage, two computerised models for predicting breakthrough UTI have been published. In both studies, high-grade VUR, a history of febrile UTI, bladder bowel dysfunction and female sex were predictors of recurrent febrile infections, which are more or less the same as for the prediction of the non-resolution of VUR. These studies do not, however, specifically deal with high-grade infant VUR but include children with all grades of VUR presented at all ages.

There are several arguments in favour of the rationale for using variables from a cross-section of about one year of age instead of data available at the time of VUR diagnosis. The first is to have complete, comparable investigations at a clear time point, in all the study patients, in order to make the data applicable and helpful for clinicians deciding on further treatment in infants with high-grade VUR. Secondly, many of the variables with a known association with VUR resolution are not known for the individual patient at the time of VUR diagnosis. Renal scans performed in the early postnatal period are known to have poor quality, due to the physiological low renal function at that time and especially when seen in conjunction with dilated upper urinary tracts. The same holds true for estimated or investigated glomerular filtration rate measurements.

Finally, recurrent infections, despite prophylaxis, are most frequent during the first year but are only known for a small fraction of patients at the time of VUR diagnosis.

Furthermore, the timing for decision regarding intervention or surveillance at the age of one year can be supported by current urological guidelines and the management in many institutions. During the first year, antibiotic prophylaxis is the recommended first-line treatment with a conservative attitude towards early anti-reflux procedures, even though some bilateral grade 5 with heavily dilated upper urinary tracts and UTI problems might require early temporary bladder drainage. If breakthrough UTI and/or pre-existing or new renal cortical damage and/or persistent high-grade VUR
at the age of 12-24 months that is unlikely to resolve; surgical intervention should be considered according to the recent urological guidelines and also supported by the treatment algorithm in the frequently used UpToDate webpage.

TABLE 4 Number of patients in the No event (VUR > 2) and Event (VUR ≤ 2) groups with regard to each scoring value (0-14) and with comparisons between scores/patient between the groups

| Variable | No event (n = 62) | Event (n = 27) | P-value |
|----------|------------------|----------------|---------|
| Score    |                  |                |         |
| 0        | 0 (0%)           | 3 (11%)        |         |
| 2        | 0 (0%)           | 3 (11%)        |         |
| 3        | 1 (2%)           | 0 (0%)         |         |
| 4        | 8 (13%)          | 10 (37%)       |         |
| 5        | 3 (5%)           | 3 (11%)        |         |
| 6        | 3 (5%)           | 2 (7%)         |         |
| 7        | 18 (30%)         | 4 (15%)        |         |
| 8        | 6 (10%)          | 0 (0%)         |         |
| 9        | 4 (7%)           | 0 (0%)         |         |
| 10       | 7 (12%)          | 1 (4%)         |         |
| 11       | 7 (12%)          | 1 (4%)         |         |
| 12       | 4 (7%)           | 0 (0%)         |         |
| Score    | 7.9 (2.8)        | 4.5 (2.7)      | <.0001  |
| Score    | 7 (3-14)         | 4 (0-11)       |         |

Note: For categorical variables, n (%) is presented. For continuous variables, the mean (SD)/median (min-max) is presented.

Scoring variable = 4*gender [0,1] + 3*breakthrough UTI [0,1] + 2*type of renal abnormality [0,1,2] + 3*GFR categorised [0,1].

TABLE 5 Probability of the spontaneous resolution/downgrading of VUR (grade ≤2) in infantile high-grade VUR, estimated from data at age one year. High probability in scores of ≤4, intermediate probability scores 5-7 and low probability scores ≥8

| Score | P     |
|-------|-------|
| High  |       |
| 0     | .91   |
| 1     | .86   |
| 2     | .79   |
| 3     | .69   |
| 4     | .57   |
| Intermediate |     |
| 5     | .44   |
| 6     | .32   |
| 7     | .22   |
| Low   |       |
| 8     | .14   |
| 9     | .09   |
| 10    | .06   |
| 11    | .03   |
| 12    | .02   |
| 13    | .01   |
| 14    | .01   |

CONCLUSION

This model provides a practical tool in the management of infants born with high-grade reflux. Scores ≥8 at one year of age indicate a high risk of persistent reflux and the need for close surveillance or surgical intervention.

CONFLICT OF INTEREST

None of the authors has any conflict of interest.

ORCID
Sofia Sjöström [https://orcid.org/0000-0002-1376-9092]

REFERENCES
1. Godley ML, Desai D, Yeung CK, Dhillon HK, Duffy PG, Ransley PG. The relationship between early renal status, and the resolution of vesico-ureteric reflux and bladder function at 16 months. BJU Int. 2001;87(6):457-462.
2. Silva JM, Diniz JS, Lima EM, Vergara RM, Oliveira EA. Predictive factors of resolution of primary vesico-ureteric reflux: a multivariate analysis. BJU Int. 2006;97(5):1063-1068.
3. Sjöström S, Sillen U, Jodal U, Sameby L, Sixt R, Stokland E. Predictive factors for resolution of congenital high grade vesicoureteral reflux in infants: results of univariate and multivariate analyses. J Urol. 2010;183(3):1177-1184.
4. Knudson MJ, Austin JC, Wald M, Makhlof AA, Niederberger CS, Cooper CS. Computational model for predicting the chance of early resolution of vesico-ureteric reflux. J Urol. 2006;176(5):1905-1911.
resolution in children with vesicoureteral reflux. J Urol. 2007;178(4 Pt 2):1824-1827.
5. Estrada CR, Passerotti CC, Graham DA, et al. Nomograms for predicting annual resolution rate of primary vesicoureteral reflux: results from 2,462 children. J Urol. 2009;182(4):1535-1541.
6. Kirsch AJ, Arlen AM, Leong T, et al. Vesicoureteral reflux index (VURx): a novel tool to predict primary reflux improvement and resolution in children less than 2 years of age. J Pediatr Urol. 2014;10(6):1249-1254.
7. Arlen AM, García-Roig M, Weiss AD, Leong T, Cooper CS, Kirsch AJ. Vesicoureteral reflux index: 2-institution analysis and validation. J Urol. 2016;195(4 Pt 2):1294-1299.
8. Hidas G, Billimek J, Nam A, et al. Predicting the risk of breakthrough urinary tract infections: primary vesicoureteral reflux. J Urol. 2015;194(5):1396-1401.
9. Arlen AM, Alexander SE, Wald M, Cooper CS. Computer model predicting breakthrough febrile urinary tract infection in children with primary vesicoureteral reflux. J Pediatr Urol. 2016;12(5):288.e1–e5.
10. Knudson MJ, Austin JC, McMillan ZM, Hawtrey CE, Cooper CS. Predictive factors of early spontaneous resolution in children with primary vesicoureteral reflux. J Urol. 2007;178(4 Pt 2):1684-1688.
11. Sjöström S, Sillen U, Bachelard M, Hansson S, Stokland E. Spontaneous resolution of high grade infantile vesicoureteral reflux. J Urol. 2004;172(2):694-698; discussion 9.
12. Nordenstrom J, Holmdahl G, Brandstrom P, et al. The Swedish infant high-grade reflux trial: study presentation and vesicoureteral reflux outcome. J Pediatr Urol. 2017;13(2):130-138.
13. Schwartz GJ, Muñoz A, Schneider MF, et al. New equations to estimate GFR in children with CKD. J Am Soc Nephrol: JASN. 2009;20(3):629-637.
14. Lebowitz RL, Olbing H, Parkkulainen KV, Smellie JM, Tamminen-Möbius TE. International system of radiographic grading of vesicoureteric reflux. International Reflux Study in Children. Pediatr Radiol. 1985;15(2):105-109.
15. Austin PF, Bauer SB, Bower W, et al. The standardization of terminology of lower urinary tract function in children and adolescents: update report from the standardization committee of the International Children’s Continence Society. Neurourol Urodyn. 2016;35(4):471-481.
16. Jansson UB, Hanson M, Hanson E, Hellstrom AL, Sillen U. Voiding pattern in healthy children 0 to 3 years old: a longitudinal study. J Urol. 2000;164(6):2050-2054.
17. Piepsz A, Colarinha P, Gordon I, et al. Guidelines for 99mTc-DMSA scintigraphy in children. Eur J Nucl Med. 2001;28(3):BP37-41.
18. Brochner-Mortensen J, Haahr J, Christoffersen J. A simple method for accurate assessment of the glomerular filtration rate in children. Scand J Clin Lab Invest. 1974;33(2):140-143.
19. Winberg J. The 24-hour true endogenous creatinine clearance in infants and children without renal disease. Acta Paediatr. 1959;48:443-452.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sjöström S, Pivodic A, Abrahamsson K, Sixt R, Stokland E, Hansson S. A scoring system for predicting downgrading and resolution of high-grade infant vesicoureteral reflux. Acta Paediatr. 2021;110:347-356. https://doi.org/10.1111/apa.15404