0749-5161/03/1903-0162
PEDIATRIC EMERGENCY CARE

Vol. 19, No. 3
Printed in U.S.A.
10.1097/01.pec.0000081238.98249.40

Validation of a decision rule identifying febrile young girls at high risk for urinary tract infection

MARC H. GORELICK, MD, MSCE, ALEJANDRO HOBERMAN, MD, DIANA KEARNEY, RN, ELLEN WALD, MD, KATHY N. SHAW, MD, MSCE

Objective: **To validate a previously published clinical decision rule to predict risk of urinary tract infection in febrile young girls.**

Methods: **We performed a retrospective case-control study at a children's hospital emergency department in a different city than that in which the original derivation study took place. Girls younger than 2 years in whom urinalysis and urine culture were performed for evaluation of fever were eligible. Cases consisted of all patients with a positive urine culture result, defined as 50,000 or more colony-forming units per milliliter of a urinary tract pathogen ($n = 98$). A random sample of patients with a negative urine culture result ($n = 114$) was also selected as controls. The clinical prediction rule included five risk factors: age younger than 12 months, white race, temperature of 39.0°C or higher, absence of any other potential source of fever, and fever for 2 days or more. The sensitivity and false-positive rate of this rule were calculated at different cutoff values.**

Results: **The overall discriminative ability of the rule, as indicated by the area under the receiver–operator characteristic curve (AUC), was similar in this validation sample (AUC = 0.72) to that in the original study (AUC = 0.76). However, in the validation sample, the presence of three or more risk factors (rather than two or more as in the original study) appeared to be the optimum cutoff to define a positive rule, which results in an indication for obtaining further diagnostic testing (sensitivity, 88% [95% CI, 79–94%]; false-positive rate, 70% [95% CI, 61–79%]).**

Conclusion: **A simple clinical decision rule previously developed to predict urinary tract infection based on five risk factors performs similarly in a different patient population.**

From the Section of Emergency Medicine, Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin (M.H. Gorelick); Divisions of General Academic Pediatrics (A. Hoberman, D. Kearney) and Allergy, Immunology, and Infectious Diseases (E. Wald), Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; and Division of Emergency Medicine, Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania (K.N. Shaw).

Supported by grant MCJ-420648 from the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration, US Department of Health and Human Services, Rockville, MD.

Address for reprints: Marc H. Gorelick, MD, MSCE, Section of Emergency Medicine, Children's Hospital of Wisconsin MS#677, 9000 W. Wisconsin Avenue, Milwaukee, WI 53226; e-mail: mgorelic@mcw.edu

Key Words: Urinary tract infection, decision rule, predictive model

INTRODUCTION

Urinary tract infection (UTI) has been found to be a relatively common cause of fever in children, with a reported prevalence among febrile young girls in an emergency department (ED) setting of 4.3 to 8.3% (1, 2). Specific symptoms of UTI are uncommon in young children, making identification of the child with UTI difficult. However, the risk of UTI is affected by various factors, including age, race, presence of another source of fever, and height and duration of fever (1, 2). A clinical decision rule based on these five risk factors was recently developed to identify febrile girls younger than 2 years with UTI (3). In the sample in which the rule was derived, the presence of two or more risk factors was found to have a sensitivity of 95%; however, the performance of the clinical decision rule has not been validated in a different patient population. We report here the use of this published predictive model in a different ED.

METHODS

In this retrospective, case-control study, patients were selected from girls aged 1 month to 2 years seen in the ED of the Children's Hospital of Pittsburgh between July 1, 1995, and June 30, 1997, from whom a urine sample was obtained by urethral catheterization, and urinalysis and urine culture were performed as part of the evaluation of fever. Ninety-eight patients with UTI, defined as urine culture with growth of 50,000 or more colony-forming units per milliliter of a urinary tract pathogen, were identified and included in the study. A random sample of 114 girls in the same age group with a negative urine culture result was also identified by selecting every fifth patient seen during the same time period who had a negative urine culture result. Medical records were reviewed, and data were abstracted regarding the five risk factors included in the predictive model: age (<12 months vs 12–23 months), race (white vs nonwhite), duration of fever (<2 days vs ≥2 days), height of fever (<39°C vs ≥39°C), and absence of another potential source of fever. The presence of another source of fever in this study was defined according to the discharge diagnosis; diagnoses such as "fever," "fever with no source," or "viral syndrome" were considered to indicate the absence of a source of fever, while patients with specific diagnoses, including otitis media, upper respiratory infection, or gastroenteritis, were considered to have another potential source of fever.

Logistic regression was performed to evaluate the association between each risk factor and the presence of UTI, adjusting for confounding by other variables. Adjusted odds ratios and 95% confidence intervals were calculated. A summary score, with each risk factor assigned one point, was calculated for each patient. The sen-

sitivity and specificity of the rule at each possible cutoff score was calculated, with 95% confidence intervals, and the results were plotted on a receiver–operator characteristic (ROC) curve. The overall ability of the rule to discriminate between patients with and without UTI is indicated by the area under the ROC curve.

RESULTS

The characteristics of the patients in this validation sample are shown in Table 1. Compared with the original derivation sample from the Children's Hospital of Philadelphia, the patients in the Children's Hospital of Pittsburgh study were more likely to be white (66% vs 12%) and to lack another source of fever (76% vs 22%). Even among those without UTI, only 38% had another possible explanation for the fever. As this is a case-control study, not a cohort study (as the original Philadelphia study was), the distribution of the risk factors under study does not necessarily reflect their distribution in the underlying source population at Children's Hospital of Pittsburgh.

Three of the five risk factors (white race, absence of another source of fever, and fever for at least 2 days) were strongly associated with UTI, as shown in Table 2, and a fourth (fever ≥39°C) had a borderline association. Only age less than 12 months was not significantly associated with UTI. The performance of the clinical decision rule based on all five risk factors is illustrated in the ROC curve (Fig. 1). The ROC curves for the derivation and validation samples are quite similar. Overall, the discriminative ability of the rule was similar in the Children's Hospital of Pittsburgh study population (area under the curve = 0.72) to that in the original Children's Hospital of Philadelphia derivation sample (area under the curve = 0.77). However, the curve for the sample in the current study is shifted to the right compared with that of the original study. At each possible cutoff, the validation sample shows a higher true-positive rate and false-positive rate. For example, the presence of two or more risk factors identified 99% of the patients with UTI in this sample (ie, sensitivity, 99% [95% CI, 94–100%]) compared with a sensitivity of 95% (95% CI, 85–99%) in the derivation study. The false-positive rate at this cutoff was 90% (95% CI, 83–95%) in the validation sample compared with 69% (95% CI, 66–72%) in the original derivation study. In the current study, the presence of three or more risk factors had a sensitivity of 88% (95% CI, 79–94%) and a false-positive rate of 70% (95% CI, 61–79%).

TABLE 1
Demographic and clinical characteristics of subjects from the validation and derivation groups

	Children's Hospital of Pittsburgh (validation group)	Children's Hospital of Philadelphia (derivation group)
Mean age (± SD), mo	9.4 ± 5.8	11.7 ± 6.7
Percent younger than 1 year	67	52
Percent white	66	11.6
Mean temperature (± SD), °C	39.4 ± 0.8	39.4 ± 0.7
Percent with temperature of 39°C or higher	74	70
Mean duration of symptoms, d	1.9 ± 2.0	1.8 ± 1.8
Percent with symptoms ≥ 2 days	38	19
Percent with another source of fever	24	78

TABLE 2
Association between clinical risk factors and urinary tract infection

Risk Factor	Children's Hospital of Pittsburgh (validation group), adjusted OR (95% CI)	Children's Hospital of Philadelphia (derivation group), adjusted OR (95% CI)
White race	3.0 (1.4–6.4)	7.5 (4.2–13.5)
Age < 12 months	0.85 (0.4–1.7)	3.0 (1.5–5.9)
Temperature > 39°C	1.6 (0.8–3.3)	2.6 (1.3–5.4)
Absence of another source of fever	13.4 (4.1–44)	2.4 (1.3–4.5)
Fever for 2 days or more	3.3 (1.6–6.9)	2.0 (1.1–3.6)

DISCUSSION

Clinical decision rules, which categorize patients into different levels of risk, may be used to guide patient management. However, rules that categorize patients accurately in one population may perform less well when applied in a different setting (4). External validation is therefore crucial in determining the general applicability of a predictive model. In this study, a previously published decision rule was found to have similar overall diagnostic accuracy as in the population in which the rule was first derived.

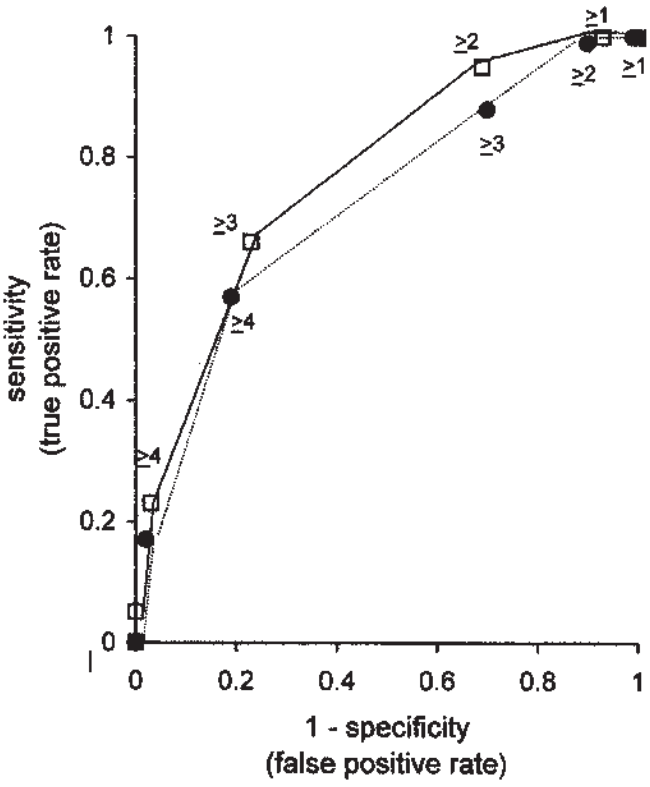

FIG. 1. Receiver–operator characteristic curves for the clinical decision rule. The curve for the validation sample (Children's Hospital of Pittsburgh) is shown with solid circles, while the derivation sample (Children's Hospital of Philadelphia) curve is shown with open squares. The cutoff value (number of risk factors present) for a positive rule is shown next to each point on the graph.

Although both studies were performed in an ED, the patient characteristics are different. Patients in the validation study were less likely, for example, to have another potential source of fever (including such nonspecific entities as gastrointestinal symptoms or upper respiratory infection). The validation population also had a higher proportion of white patients when compared with the sample from Philadelphia. It is important to note that the patients in this study were drawn from those in whom urinalysis and urine culture were performed as part of the evaluation for fever. Because testing for UTI was not performed universally among young febrile girls in the ED, we were unable to determine the characteristics of all potentially eligible patients. However, in a previous study from the Children's Hospital of Pittsburgh, 56% of febrile young children evaluated in the ED were white (1). These results suggest that the clinical decision rule performs similarly in fairly diverse patient populations.

The retrospective nature of this study is an important potential limitation. Specifically, although we do not have data on the patients in whom urine testing was not performed, our experience is that clinicians in the ED at Children's Hospital of Pittsburgh have been influenced by some of the risk factors included here (especially race, absence of another source of fever, and duration of fever) since the original study of prevalence of UTI was conducted (1) to determine whether or not to perform such tests. In other words, the decision to perform the gold standard test, in this case urine culture, is influenced by the results of the screening test (ie, the result of the clinical prediction rule). This may produce a type of bias called spectrum bias, or test referral bias, leading to a study population in which each of these risk factors is overrepresented when compared with the underlying source population. (Because not all eligible subjects have their disease status verified with the gold standard, this is sometimes also referred to as verification bias (5).) This would result in overestimation of both the apparent true-positive rate (ie, sensitivity) and false-positive rate (6, 7). Indeed, this was noted in the present study, with both true- and false-positive rates at each cutoff being higher in the validation sample than in the derivation sample. If the proportion of all febrile young children at the Pittsburgh ED in whom urine testing was performed was known, we could attempt to estimate the degree of verification bias and correct for it (5). We do not have precise estimate of the rate of urine testing in the target population. However, the authors' experience is that such testing is performed fairly liberally.

The choice of a cutoff value is determined by the relative importance of sensitivity and specificity, which, in turn, will depend on the practice setting and the availability of follow-up. The results of this study suggest that obtaining a urine culture in all girls with two or more risk factors would lead to the identification of virtually all cases of UTI but would eliminate only 10% of negative test results. Limiting urine culture to those girls with three or more of the clinical features would still have acceptable sensitivity (88%) but would have the effect of further decreasing unnecessary tests by about a third. However, as noted previously, the sensitivity and specificity are likely underestimates of what would be expected in an unselected population. Accordingly, these results must be interpreted with caution. In the ED, where sensitivity is typically more highly valued, it may be prudent to use a more inclusive definition of *high risk* (ie, two or more risk factors present) to avoid missing patients with UTI. In a practice setting, where close follow-up is ensured, the more restrictive definition of a positive rule (ie, at least three risk factors present) may be acceptable.

CONCLUSIONS

Recognition of UTI as a relatively common source of fever in young children has led to recommendations for increased vigilance in testing for UTI among such patients (8). Because signs and symptoms of UTI in such children are nonspecific, a high index of suspicion is necessary. A selective approach, based on clinical risk factors, has been proposed to identify children at sufficiently low risk of UTI, in whom diagnostic testing (ie, urinalysis and urine culture) may be safely deferred, potentially decreasing unnecessary testing and treatment (3). We have shown that the previously reported clinical risk factors were also predictive of UTI in this population, and that the ability of the clinical decision rule to discriminate between low- and high-risk patients was similar in a different patient population than in the one in which it was derived. Further study, particularly in more diverse practice settings, can help define the effectiveness and efficiency of a rule-based approach to testing for UTI in febrile young girls.

REFERENCES

1. Hoberman A, Han-Pu C, Keller DM, et al. Prevalence of urinary tract infection in febrile infants. *J Pediatr*. 1993;123:17–23.
2. Shaw KN, Gorelick MH, McGowan KL, et al. Prevalence of UTI in febrile young children in the emergency department. *Pediatrics*. 1998; 102:e15.
3. Gorelick MH, Shaw KN. Clinical decision rule identify young febrile children at risk for UTI. *Arch Pediatr Adolesc Med*. 2000;154:386–390.
4. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA*. 1997;277: 488–494.
5. Zhou XH. Correcting for verification bias in studies of a diagnostic test's accuracy. *Stat Methods Med Res*. 1998;7:337–353.
6. Sox HC, Blatt MA, Higgins MC, et al. *Medical Decision Making*. Boston, MA: Butterworth-Heinemann; 1988:116–123.
7. Gray R, Begg CB, Greenes RA. Construction of receiver operating characteristic curves when disease verification is subject to selection bias. *Med Decis Making*. 1984;4:151–164.
8. American Academy of Pediatrics. Practice parameter: the diagnosis, treatment, and evaluation of the initial urinary tract infection in febrile infants and young children. *Pediatrics*. 1999;103:843–852.