Validation of a Clinical Prediction Rule for the Differentiation Between Septic Arthritis and Transient Synovitis of the Hip in Children

By Mininder S. Kocher, MD, MPH, Rahul Mandiga, BS, David Zurakowski, PhD, Carol Barnewolt, MD, and James R. Kasser, MD

Investigation performed at the Departments of Orthopaedic Surgery, Biostatistics, and Radiology, Children's Hospital, Boston, Massachusetts

Background: The differentiation between septic arthritis and transient synovitis of the hip in children can be difficult. The purpose of the present study was to validate a previously published clinical prediction rule for this differentiation in a new patient population.

Methods: We prospectively studied children who presented to a major children’s hospital between 1997 and 2002 with an acutely irritable hip. As in the previous study, diagnoses of septic arthritis (fifty-one patients) and transient synovitis (103 patients) were operationally defined on the basis of the white blood-cell count in the joint fluid, the results of cultures of joint fluid and blood, and the clinical course. Univariate analysis and multiple logistic regression were used to compare the two groups. The predicted probability of septic arthritis of the hip from the prediction rule was compared with actual distributions in the current patient population. The area under the receiver operating characteristic curve was determined.

Results: The same four independent predictors of septic arthritis of the hip (a history of fever, non-weight-bearing, an erythrocyte sedimentation rate of 40 mm/hr, and a serum white blood-cell count of >12,000 cells/mm$^3$ (>12.0 × 10$^9$/L)) were identified in the current patient population. The predicted probability of septic arthritis of the hip from the prediction rule was similar to the actual distributions in the current patient population. The area under the receiver operating characteristic curve for the current patient population was 0.86, compared with 0.96 in the original population.

Conclusions: Clinical prediction rules typically demonstrate diminished performance in a new patient population because they are optimally modeled to the original data set. The previously published clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children demonstrated diminished, but nevertheless very good, diagnostic performance in a new patient population.

Level of Evidence:Diagnostic study, Level I-1 (testing of previously developed diagnostic criteria in series of consecutive patients [with universally applied reference “gold” standard]). See Instructions to Authors for a complete description of levels of evidence.

Differentiation between septic arthritis and transient synovitis in a child with an acutely irritable hip is essential because the two clinical entities have vastly different treatments and potentials for sequelae$^{1,22}$. Furthermore, the early, accurate diagnosis of septic arthritis of the hip in children is critical because poor outcomes have been associated with a delay in diagnosis$^{30,31}$. Nevertheless, this differentiation between septic arthritis and transient synovitis can be difficult because both conditions often present similarly as an atraumatic, acutely irritable hip in a child who has...
progressive symptoms and signs of fever, who has a limp or refuses to bear weight, and who has limited motion, joint effusion, and abnormal laboratory findings on evaluation of the blood and the joint fluid.\(^{22,28}\)

In a previous report,\(^7\) we described a clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip that was based on four independent, multivariate predictors of septic arthritis of the hip: a history of fever, non-weight-bearing, an erythrocyte sedimentation rate of 40 mm/hr, and a serum white blood-cell count of \(>12,000 \text{ cells/mm}^3\) (\(>12.0 \times 10^9/L\)). Derived from the data on eighty-two patients with septic arthritis of the hip and eighty-six patients with transient synovitis who were evaluated between 1979 and 1996, this prediction rule demonstrated excellent diagnostic performance, with an area under the receiver operating characteristic curve of 0.96 and a significant result on the Pearson chi-square test for trend (\(p < 0.0001\)).

Clinical prediction rules typically demonstrate diminished performance in a new patient population because they are optimally modeled to the original data set.\(^{22,29}\) Therefore, the validation of a clinical prediction rule in a new patient population is essential. The purpose of the present study was to validate the previously described clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children in a new patient population.

### Materials and Methods

Institutional review board approval was obtained for this prospective study. All patients who presented to a major tertiary-care children's hospital between 1997 and 2002 with an acutely irritable hip and a differential diagnosis of transient synovitis or septic arthritis were managed on the basis of the previously published clinical practice guideline.\(^{25}\)

As in the previous study,\(^7\) the diagnoses of septic arthritis and transient synovitis were operationally defined on the basis of the white blood-cell count in the joint fluid, the results of cultures of joint fluid and blood, and the clinical course. The diagnosis of true septic arthritis was assigned when the patient had either a positive finding on culture of joint fluid or a white blood-cell count in the joint fluid of \(\geq 50,000 \text{ cells/mm}^3\) (\(\geq 50.0 \times 10^9/L\)) with a positive finding on blood culture. The diagnosis of presumed septic arthritis was explicitly assigned when the patient had a white blood-cell count in the joint fluid of \(\geq 50,000 \text{ cells/mm}^3\) (\(\geq 50.0 \times 10^9/L\)) with negative findings on cultures of the joint aspirate and blood. Thus, the group with septic arthritis included both the group with true septic arthritis and the group with presumed septic arthritis. The diagnosis of transient synovitis was explicitly assigned when the patient had a white blood-cell count in the joint fluid of \(<50,000 \text{ cells/mm}^3\) (\(<50.0 \times 10^9/L\)) with negative findings on culture, resolution of symptoms without antimicrobial therapy, and no further development of a disease process as documented in the medical record. The mean duration of follow-up was 11.8 months (range, 5.9 to 23.7 months).

Of the 213 eligible consecutive patients who were evaluated during the study period, twenty-four were diagnosed with true septic arthritis, twenty-seven were diagnosed with presumed septic arthritis, 103 were diagnosed with transient synovitis, and fifty-nine were excluded. The exclusion criteria were the same as those described in the previous study.\(^7\) The fifty-nine excluded patients included individuals in atypical groups, such as those with immunocompromise (ten patients), renal failure (three), neonatal sepsis (three), postoperative infection of the hip (two), later development of rheumatologic disease (one), or later development of Legg-Calvé-Perthes disease (one). Two patients with septic arthritis and associated proximal femoral osteomyelitis who had had symptoms for more than two weeks and who had proximal femoral radiolucency on radiographs, an intrasosseous abscess that was confirmed at the time of arthrotomy and femoral neck drilling, and bacteremia were excluded because they were not believed to pose the typical diagnostic dilemma between septic arthritis and transient synovitis. To avoid information bias associated with incomplete data analysis and to avoid selection bias associated with the inclusion of patients with presumptive and inconsistent diagnoses, a patient was excluded if joint fluid had not been obtained (or insufficient fluid had been obtained) for a cell count, gram stain, or culture (twenty-seven patients); if peripheral blood had not been obtained for a cell count or culture (five); if the white blood-cell count in the joint fluid was \(<50,000 \text{ cells/mm}^3\) (\(<50.0 \times 10^9/L\)) with negative cultures but the patient was treated with arthrotomy and intravenous administration of antibiotics (four); or if the white blood-cell count in the joint fluid was \(<50,000 \text{ cells/mm}^3\) (\(<50.0 \times 10^9/L\)) with negative cultures but the patient was managed with intravenous administration of antibiotics alone on the pediatric service (one). Of the twenty-seven patients who were excluded because of the lack of joint-fluid analysis, twenty-five had an insufficient quantity of aspirate available for both a cell count and gram stain with culture, and therefore only one test had been performed.

Data were obtained for all patients on age, gender, date of presentation, duration of symptoms, history of fever, history of chills, weight-bearing status, history of trauma, history of concurrent or recent infection, history of recent antibiotic use, temperature, erythrocyte sedimentation rate, serum white blood-cell count and differential, platelet count, hematocrit, results of blood culture, evidence of hip joint effusion on radiographs, and results of gram-staining, cell count, differential, and culture of joint fluid. C-reactive protein values were not obtained for all patients because, during the study period at our institution, the availability of C-reactive protein testing evolved from weekly testing to once-daily testing to routine testing. A history of fever was operationally defined as an oral temperature of \(>38.5^\circ C\) during the week before presentation. A history of chills was operationally defined as positive documentation of chills; a negative history of chills was coded either for documentation of no chills or for no documentation of chills. Weight-bearing status was determined on the basis of the clinical history and was considered the inability or refusal to bear weight even with support. An effusion was defined as a
side-to-side distance of >2 mm from the medial part of the femoral head to the medial part of the acetabulum on an an­teroposterior pelvic radiograph.

Univariate analysis was performed with use of the two­sample Student t test for continuous variables and with use of the Fisher exact test for categorical variables. Comparisons were made between the group with true septic arthritis and the group with presumed septic arthritis and between the group with septic arthritis and the group with transient synovitis. Stepwise multiple logistic regression with use of backward selection was performed to identify independent clinical predictors, and comparisons were made between the septic arthritis and transient synovitis groups. Variables associated with a p value of <0.20 in the univariate analysis were chosen as candidates for the multivariate model, with significance determined by the likelihood ratio chi-square test. Regression model fit was estimated with the Hosmer-Lemeshow goodness-of-fit test. Adjusted odds ratios and 95% confidence intervals were derived with the method of maximum likelihood.

A receiver operating characteristic curve was constructed to assess the diagnostic performance of the group of multivariate predictors in identifying septic arthritis. Statistical analysis was performed with use of the SPSS (version 11.0; SPSS, Chicago, Illinois) and SAS (version 6.12; SAS Institute, Cary, North Carolina) software packages.

To evaluate the diagnostic performance of the previously described clinical prediction in the current patient population, the multivariate predictors from the current patient population were compared with those from the original population. In addition, the predicted probability of septic arthritis from the prediction rule was compared with actual distributions in the current patient population. The area under the receiver operating characteristic curve for the current patient population was compared with that for the original population. The receiver operating characteristic curve is a graphic analytical technique that is used to evaluate the diagnostic performance of a test or a prediction rule. Sensitivity is plotted on the y axis, and the false-positive rate (1 – specificity) is plotted on the x axis. The area under the receiver operating characteristic curve is a summary measure of the diagnostic performance of the prediction rule. A perfect rule would approximate the upper left corner of the graph with an area under the curve of 1.0. Random guessing would be a straight line graph with an area under the curve of 0.5.

Results

**Descriptive Data**

Of the fifty-one patients with septic arthritis, twenty-four (47%) had positive results on culture and twenty-seven (53%) had negative results on culture. Of the twenty-four patients with positive results, sixteen had a positive joint-fluid culture and a positive blood culture, six had a positive joint-fluid culture and a negative blood culture, and two had a negative joint-fluid culture and a positive blood culture. Organisms isolated on culture included *Staphylococcus aureus* (fifteen patients), *Streptococcus pneumoniae* (six), *Neisseria meningitidis* (two), and group-A Streptococcus (one). Of the twenty-two patients with a positive culture of joint fluid, seventeen had positive gram stains of joint fluid. There were no positive gram stains of joint fluid from the patients who had negative joint-fluid cultures.

**Univariate Analysis: True Septic Arthritis Compared with Presumed Septic Arthritis**

The twenty-four patients who had true septic arthritis differed significantly from the twenty-seven patients who had presumed septic arthritis with regard to a history of recent antibiotic use (two patients [8%] compared with ten patients [37%], p < 0.001), a history of chills (six patients [25%] compared with zero patients [0%], p < 0.001), temperature elevation (38.1°C ± 0.8°C compared with 37.4°C ± 0.9°C, p = 0.02), erythrocyte sedimentation rate (52.9 ± 20.1 compared with 41.1 ± 19.9 mm/hr, p = 0.04), a history of fever (twenty-three patients [96%] compared with nineteen patients [70%], p = 0.005), male gender (fifteen patients [63%] compared with seven patients [26%], p = 0.03), and serum white blood-cell differential (bands [8.2% ± 9.4% compared with 3.8% ± 2.5%, p = 0.02). There were no significant differences (p > 0.05) with regard to age, duration of symptoms, effusion on radiographs, non-weight-bearing status, hematocrit, platelet count, serum white blood-cell count, or serum white blood-cell differential for neutrophils, lymphocytes, monocytes, atypical lymphocytes, eosinophils, and basophils.

**Univariate Analysis: Septic Arthritis Compared with Transient Synovitis**

The patients who had septic arthritis differed significantly from those who had transient synovitis with regard to gender,

<table>
<thead>
<tr>
<th>Multivariate Predictor</th>
<th>Regression Coefficient</th>
<th>Likelihood Ratio Test</th>
<th>P Value</th>
<th>Adjusted Odds Ratio</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of fever</td>
<td>1.5</td>
<td>11.6</td>
<td>&lt;0.001</td>
<td>4.4</td>
<td>1.8-10.4</td>
</tr>
<tr>
<td>Non-weight-bearing</td>
<td>1.8</td>
<td>14.2</td>
<td>&lt;0.001</td>
<td>5.9</td>
<td>2.2-16.1</td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate ≥40 mm/hr</td>
<td>1.5</td>
<td>11.2</td>
<td>&lt;0.001</td>
<td>4.5</td>
<td>1.8-10.9</td>
</tr>
<tr>
<td>Serum white blood-cell count</td>
<td>1.4</td>
<td>10.7</td>
<td>&lt;0.001</td>
<td>4.1</td>
<td>1.7-10.0</td>
</tr>
</tbody>
</table>
Multivariate Analysis: Septic Arthritis Compared with Transient Synovitis

We identified the same four independent multivariate predictors of septic arthritis in the current population as we had in the derivation population: a history of fever, non-weight-bearing status, an erythrocyte sedimentation rate of ≥40 mm/hr, and a serum white blood-cell count of >12,000 cells/mm³ (>12.0 × 10⁹/L) (Table I). The Hosmer-Lemeshow goodness-of-fit test revealed no significant departure from good model fit (p > 0.05).

Algorithm for Probability of Septic Arthritis

The actual distribution of septic arthritis in the current population was similar to the predicted probability of septic arthritis derived from the original population for the algorithm based on all sixteen combinations of the four predictors and for the simplified algorithm based on the number of predictors (see Appendix).

For a patient with zero predictors, the predicted probability of septic arthritis from the previous study was <0.2% and the actual distribution in the current study was 2.0%. For a patient with one predictor, the predicted probability of septic arthritis from the previous study was 3.0% and the actual distribution in the current study was 9.5%. For a patient with two predictors, the predicted probability of septic arthritis from the previous study was 40.0% and the actual distribution in the current study was 35.0%. For a patient with three predictors, the predicted probability of septic arthritis from the previous study was 93.1% and the actual distribution in the current study was 72.8%. For a patient with four predictors, the predicted probability of septic arthritis from the previous study was 99.6% and the actual distribution in the current study was 93.0%.

Receiver Operating Characteristic Curves

Receiver operating characteristic curves for the clinical prediction rule in the original population and the current population are shown in Figure 1. The data values for the original population and the current population are shown in Table II. The area under the curve for the original population was 0.96, indicating that this group of four multivariate predictors demonstrated excellent diagnostic performance in the identification of septic arthritis. The area under the curve for the current patient population was 0.86, indicating very good diagnostic performance.

Discussion

Clinical prediction rules are intended to make the art of diagnosis more objective by allowing the clinician to estimate the probability of a diagnostic outcome and to classify patients according to the risk of disease. Prediction rules originally took the form of clinical aphorisms based on the empiric experience of senior clinicians; however, more recently, they have been derived from evidence-based mathematical analyses. For the clinician who is faced with the important but often difficult task of differentiating between septic arthritis and transient synovitis of the hip in children, the clinical prediction rule that we previously described may be useful for guiding the diagnostic workup and establishing a timely and accurate diagnosis.
Multivariate predictors of septic arthritis that had been identified in the original population were found in the current population. The predicted probability of septic arthritis of the hip from the prediction rule was similar to the actual distributions in the current patient population. The area under the receiver operating characteristic curve for the current patient population was 0.86, indicating that this group of four multivariate predictors had very good diagnostic performance for identifying septic arthritis.

Jung and colleagues performed a study, similar to our original derivation study, of ninety-seven children with transient synovitis of the hip and twenty-seven children with septic arthritis of the hip. They identified differences between the two groups with use of univariate analysis and identified five independent multivariate predictors of septic arthritis: a temperature of >37°C, an erythrocyte sedimentation rate of >20 mm/hr, a C-reactive protein level of >1.0 mg/dL (>0.10 mg/L), a serum white blood-cell count of >11,000/mm³ (>11.0 x 10⁹/L), and a joint-space difference of >2 mm on radiographs. They developed a different algorithm that was based on the thirty-two combinations of the five predictors. The area under the receiver operating characteristic curve for their prediction rule was 0.986.

The limitations of the present study and the original derivation study include the lack of C-reactive protein data. C-reactive protein has been shown to have greater benefit than the erythrocyte sedimentation rate for the diagnosis of septic arthritis in children. However, during the study period at our institution, the availability of C-reactive protein testing evolved from weekly testing to once-daily testing to routine testing. Thus, the results of pretreatment C-reactive protein testing were available for only 43% (twenty-two) of the fifty-one patients with septic arthritis and 39% (forty) of the 103 patients with transient synovitis in the present study. In order to avoid biases associated with incomplete data analysis and patient selection, C-reactive protein data were not incorporated into the analysis of the prediction rule.

As in the previous study, approximately half of the patients who had septic arthritis in the present study had negative cultures and were considered to have “presumed septic arthritis.” This group with presumed septic arthritis had symptoms similar to those associated with true septic arthritis, negative cultures, and a high white blood-cell count in the joint fluid (≥50,000 cells/mm³ [≥50.0 x 10⁹/L]). As in the previous study, the patients who had true septic arthritis in the present study appeared to be more sick with fever, chills, an elevated erythrocyte sedimentation rate, and an altered white blood-cell count differential. In addition, the group
with presumed septic arthritis had a significantly higher percentage of patients with a history of recent antibiotic use as well as a significantly higher percentage of female patients. It is unclear what this presumed septic arthritis actually represented: partially treated septic arthritis, bacterial arthritis with organisms that were difficult to grow on culture, viral arthritis, arthritis resulting from atypical organisms, inflammatory or rheumatic arthritis, trauma, periarticular osteomyelitis, or an autoimmune process. Nevertheless, these patients are typically treated identically, with urgent surgical drainage and antibiotics10.

In conclusion, the previously developed clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children demonstrated diminished, but nevertheless very good, diagnostic performance in the current patient population. Further research is needed to examine its performance in new clinical settings and geographic locations. As our institution is a tertiary-care children’s hospital, differentiating variables may vary in a community hospital setting. The clinical prediction rule was developed to differentiate between septic arthritis and transient synovitis in the typical clinical scenario. Atypical patients, such as those with immunocompromise, neonatal sepsis, renal failure, postoperative septic arthritis, and associated proximal femoral osteomyelitis were excluded. Thus, the prediction rule may not be applicable in these clinical settings. Finally, it should be emphasized that a clinical prediction rule is not meant to be used as a rigid guideline or to replace clinical judgment. The goal of this clinical prediction rule is to aid in the often vexing differentiation between septic arthritis and transient synovitis of the hip in children by stratifying patients according to the risk of septic arthritis. Clinical judgment is still necessary in the further management of these patients. Patients with a minimal probability of septic arthritis (zero predictors) in the appropriate clinical setting with close follow-up may be managed with observation only, whereas patients with a high probability of septic arthritis (four predictors) may be candidates for aspiration in the operating room instead of the radiology suite given the greater likelihood that they will require surgical drainage.

Appendix

Tables showing the results of univariate analysis and the probability of septic arthritis are available with the electronic versions of this article, on our web site at www.jbjs.org (go to the article citation and click on “Supplementary Material”) and on our quarterly CD-ROM (call our subscription department, at 781-449-9780, to order the CD-ROM).

References


