Pharyngitis accounts for about 6% of visits by children to primary care physicians each year in high-income nations. Group A streptococcus is found in 30%–40% of cases of childhood pharyngitis; the remaining cases are considered viral. Antibiotic treatment is indicated for group A streptococcal infection to prevent suppurative (e.g., retropharyngeal abscess and quinsy) and nonsuppurative complications (e.g., acute rheumatic fever and rheumatic heart disease) and to reduce the duration of symptoms and the spread of the condition. In settings where poststreptococcal diseases have become uncommon, such as Western Europe and North America, the public health goal is shifting from preventing complications to minimizing the inappropriate use of antibiotic agents to contain antimicrobial resistance. However, 60%–70% of the visits by children with pharyngitis to American primary care physicians result in antibiotic agents being prescribed.

Because signs and symptoms of streptococcal and viral pharyngitis overlap, most experts recommend that the diagnosis of group A streptococcal infection be confirmed by a throat culture or rapid antigen detection test. Whereas European guidelines suggest all children with pharyngitis undergo such testing, North American guidelines recommend that clinicians select...
patients on the basis of clinical and epidemiologic grounds.\textsuperscript{11–13} Currently, there is no guidance from the Canadian Medical Association or Canadian Paediatric Society for the management of pharyngitis.

Various clinical prediction rules that combine signs and symptoms have been proposed to help clinicians define groups of patients according to the clinical likelihood of group A streptococcal infection.\textsuperscript{15–18} These rules aim to identify patients at low risk in whom the disease can be managed without further testing and without antibiotic treatment, and patients at high risk who could receive empiric antibiotic treatment without testing.\textsuperscript{16} Clinical prediction rules for pharyngitis have not been sufficiently validated for clinical practice and have never been compared head-to-head in a single pediatric population from a high-income country.\textsuperscript{18}

The purpose of our study was to externally validate and directly compare the diagnostic accuracy of relevant rules-based selective testing strategies with original data from a French prospective multicentre cohort of children with pharyngitis. To optimize this validation study, we first conducted a systematic review of existing clinical prediction rules.

\textbf{Methods}

\textbf{Systematic review}

All published studies that reported the derivation or update of clinical prediction rules for group A streptococcal pharyngitis were eligible for inclusion. We defined a clinical prediction rule as a decision-making tool that included 2 or more variables obtained from the history, physical examination or simple diagnostic tests and provided a probability of a disease or suggested a clinical course of action for an individual patient.\textsuperscript{19–23} All studies that provided data specific to children were eligible. For clinical prediction rules updated by their original authors, only the updated study was included. We excluded clinical prediction rules that were explicitly intended for use in low-resource settings (i.e., those derived in developing countries explicitly intended for use in settings where rapid antigen detection tests are not routinely available). We systematically searched MEDLINE and Embase for studies using the search strategy described in Appendix 1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). We screened the titles and abstracts of the identified studies; if they were considered potentially relevant (or in cases of uncertainty), we evaluated the full text.

We extracted a prespecified set of data related to study setting and design, and the characteristics of patients and clinical prediction rules. We assessed the quality of included studies using a combination of items from methodological standards for clinical prediction rules\textsuperscript{19–22} and clinical prediction models,\textsuperscript{24–26} and from QUADAS-2.\textsuperscript{27} (Appendix 2, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). If published reports referred to previous papers, we extracted the corresponding items. We classified clinical prediction rules according to the hierarchy of evidence for clinical prediction rules\textsuperscript{21} (with the assessment of validation based on results from previous reviews).\textsuperscript{15–18} One reviewer (JFC) performed the study selection, data extraction, and assessment of study quality and level of evidence, which were then checked by a second reviewer (MC). Discrepancies were discussed and resolved by consensus.

\textbf{External validation cohort}

We used data from a study for which patients and methods were described previously.\textsuperscript{28–30} Briefly, this French prospective multicentre study aimed to evaluate the diagnostic accuracy of a rapid antigen detection test in children aged 3–14 years, with throat culture as the clinical reference standard. Seventeen private office-based pediatricians participated. From Oct. 1, 2010, to May 31, 2011, double throat swabs were collected from 678 consecutive children who had a diagnosis of pharyngitis and did not receive antibiotics for 7 days before inclusion. Swab 1 was used for the rapid antigen detection test (StreptAtest, Dectrapharm, France) and swab 2 was sent to the hospital laboratory for throat culture on a blood agar plate according to standard methods.

The following clinical features were recorded for each patient: age, sex, sudden onset of sore throat, maximum body temperature (as reported by a parent), throat pain, cough, rhinorrhea, conjunctivitis, headache, erythema of the pharynx, tonsillar swelling, tonsillar exudate, palatal petechiae, nausea, abdominal pain, vomiting, diarrhea, presence of anterior cervical lymph nodes, size of the nodes (< 1, 1–2 or > 2 cm), tenderness of the nodes, skin rash, and aspect of the rash (scarlatiniform or nonspecific).

Microbiologists were blinded to individual clinical data and results of the rapid antigen detection tests. Parent and patient approval for participation was obtained before inclusion. The study protocol was approved by the Ile-de-France XI institutional review board (no. 09016).

\textbf{Statistical analysis}

Our primary outcome was the diagnostic accuracy of rules-based selective testing strategies,
with throat culture as the reference standard. After reviewing the literature and discussing until consensus within the review team, and assuming a prevalence of group A streptococcal infection of 35% and a maximally acceptable antibiotics prescription rate of 40%, we defined the target zone of accuracy as sensitivity and specificity greater than 85% (Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1).2,21–33 For each rules-based selective testing strategy, we used a graphical approach to test whether the one-sided rectangular 95% confidence region for sensitivity and specificity lay entirely within the target zone of accuracy.34

Our secondary outcome was the number of rapid antigen detection tests used after each clinical prediction rule. A selective testing strategy was defined as being clinically relevant if less than 80% of children would undergo testing after its application (experts’ opinions). For each selective testing strategy, we tested whether the number of tests used was less than 80% using a one-sided exact binomial probability test.

Each clinical prediction rule was applied to each patient from the validation cohort as follows: low risk of group A streptococcal infection, no further testing or antibiotic therapy; intermediate risk of infection, rapid antigen detection testing for all patients and antibiotic therapy for those with positive test results; high risk of infection, empiric antibiotic treatment. The result of the selective testing strategy was considered positive if the patient would have received antibiotic treatment and negative if the patient would not have received antibiotic treatment after the rule’s application. For rules that rely on scoring systems, we chose cut-offs for defining risk groups according to previously published studies, when available (Appendix 4, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). This method allowed us to calculate sensitivity, specificity, C-index and number of tests used for each rules-based strategy. We also performed receiver operating characteristic (ROC) analysis for each scoring system.

Another secondary objective was to assess the calibration of clinical prediction rules. For each risk level defined by the rule (low v. intermediate v. high risk of group A streptococcal infection), we compared the probability of group A streptococcal infection seen in our validation cohort (observed probability) to that reported by the authors of the clinical prediction rule in their cohort (predicted probability). We adjusted predicted probabilities for the prevalence of group A streptococcus in the validation cohort.22,35 We evaluated calibration using plots and by assessing whether clinical prediction rules were able to identify patients at low risk (probability ≤ 12%, mean prevalence of asymptomatic carriage of group A streptococcus in children) or high risk (probability ≥ 85%, minimal predictive value of a positive result on rapid antigen detection testing).

Values were missing for 0.4%–6.2% patients for each clinical variable. We used multiple imputations with chained equations (m = 10) with predictive mean matching for continuous variables and logistic regression for categorical variables to generate values for missing data (Appendix 5, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). We used Stata/SE 13 (StataCorp, College Station, Texas) for data analysis.

**Results**

**Systematic review**

We performed our literature search on Aug.21, 2014. From 1272 articles identified, we assessed the full text of 38 reports and included 8 articles in our final total (Figure 1). The studies were published between 1975 and 2010 and involved a variety of settings, with a median of 313 patients (range 94–3877) per study, and a median prevalence of group A streptococcus of 36% (range 21%–48%) (Table 1). Eight clinical prediction rules were identified for external validation (Table 2).36–41 The median number of predictors in each rule was 9 (range 5–14) (Appendix 6, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). The 5 most frequent predictors used were tenderness of lymph nodes, fever, age, tonsillar swelling and exudate.

The quality of included studies is summarized in Appendix 7 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). All of the studies were prospective. The reference standard used and inclusion criteria were sufficiently reported for 5 of 8 studies (63%). Sensitivity and specificity of the clinical prediction rule was reported in 4 of the studies, with corresponding confidence intervals in 2. There are several external validation studies for the Breese39 McIsaac40 and Wald42 scores. Attia’s clinical prediction rule43 was validated once (Appendix 7). The remaining 4 clinical prediction rules56–58,41 have never been validated. The statistical analysis used for the derivation of clinical prediction rules was poor. All of the models relied on categorization of continuous predictors, such as age and fever. Only 3 of 8 reports (38%) described the use of multivariable analysis. Only 1 article described interaction testing. Assessment of discrimination
and calibration of the model were never reported.

**External validation cohort**
A total of 678 children met our inclusion criteria; we excluded 1 patient with an uninterpretable rapid antigen detection test result and 1 patient whose throat swab was lost. Thus, our analysis included data from 676 children (313 girls [46%], mean age 6.1 [± 2.5] yr). Prevalence of group A streptococcus was 41% (95% confidence interval [CI] 38–45), and rapid antigen detection test sensitivity and specificity were 93% (95% CI 89–95) and 88% (95% CI 85–91), respectively. We had to adapt some of the clinical prediction rules because the authors provided insufficient detail for validation or some clinical variables were not assessed in the validation set; 2 rules could not be validated (Appendix 4).36,41

**Diagnostic accuracy of selective testing strategies**
The sensitivity of rules-based selective testing strategies ranged from 66% (95% CI 61–72)38 to 94% (95% CI 92–97)40; specificity ranged from 40% (95% CI 35–45)37 to 88% (95% CI 85–91)43; and the C-index ranged from 0.64 (95% CI 0.61–0.67)37 to 0.87 (95% CI 0.85–0.90)43 (Table 3, Figure 2). None of the rules-based strategies fit within our target zone of accuracy. The area under the ROC curve of clinical scores ranged from 0.56 (95% CI 0.52–0.60)40 to 0.62 (95% CI 0.58–0.66)43.

---

**Figure 1:** Selection of clinical prediction rules included in the study.

**Rapid antigen detection testing**

Use of rapid antigen detection tests after applying the clinical prediction rule ranged from 24% (95% CI 21–27)\textsuperscript{37} to 86% (95% CI 84–89)\textsuperscript{43} (Table 3). Further exploratory analyses found a strong correlation between the number of tests used after the clinical prediction rule and the accuracy of the selective testing strategy as measured by the C-index (Spearman’s \( \rho = 0.94, p < 0.01 \); Appendix 9, available at www.cmaj.ca/lookup/suppl doi:10.1503/cmaj.140772/-/DC1).

**Calibration of clinical prediction rules**

The prevalence of group A streptococcus increased with increasing risk level as defined by each of the clinical prediction rules. The calibration of 3 clinical prediction rules could not be assessed because the authors did not report prevalence by risk level.\textsuperscript{36,38,41} For the remaining rules, calibration was poor (Appendices 10 and 11, available at www.cmaj.ca/lookup/suppl doi:10.1503/cmaj.140772/-/DC1). None of the clinical prediction rules were able to identify patients at low risk (probability of group A streptococcus \( \leq 12% \)) or high risk (probability \( \geq 85% \)). The highest observed probability of group A streptococcus was 81% (score \( \geq 4 \) based on the clinical

---

**Table 1: Description of included studies**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Study period</th>
<th>Country</th>
<th>Setting</th>
<th>Aim</th>
<th>Study design</th>
<th>Reference test</th>
<th>Presenting signs and symptoms</th>
<th>Age range, yr</th>
<th>n</th>
<th>Group A streptococcus, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forsyth\textsuperscript{36}</td>
<td>Not reported</td>
<td>USA</td>
<td>Unclear</td>
<td>Derivation</td>
<td>Prospective multicentre</td>
<td>Throat culture</td>
<td>Unsolicited complaint of sore throat &lt; 1 wk duration</td>
<td>( \leq 14^* )</td>
<td>234</td>
<td>30.3*</td>
</tr>
<tr>
<td>Breese\textsuperscript{39}</td>
<td>1973–1975</td>
<td>USA</td>
<td>Office-based</td>
<td>Update, external validation</td>
<td>Prospective multicentre</td>
<td>Throat culture</td>
<td>Acute respiratory illness</td>
<td>Unclear</td>
<td>3877</td>
<td></td>
</tr>
<tr>
<td>Fujikawa \textit{et al.}\textsuperscript{41}</td>
<td>1982–1983</td>
<td>Japan</td>
<td>Pediatrics department</td>
<td>Derivation</td>
<td>Prospective single centre</td>
<td>Throat culture</td>
<td>Acute upper respiratory infection and suspicion of bacterial infection</td>
<td>Unclear</td>
<td>271</td>
<td>39.9</td>
</tr>
<tr>
<td>Wald \textit{et al.}\textsuperscript{42}</td>
<td>1990–1992</td>
<td>USA</td>
<td>Emergency department and walk-in clinic</td>
<td>Derivation</td>
<td>Prospective single centre</td>
<td>Throat culture</td>
<td>Acute onset of sore throat (&lt; 48 h duration), history of fever (tactile) or a documented oral temperature ( \geq 38.3^\circ C ) at presentation or within the preceding 24 h</td>
<td>2–16</td>
<td>365</td>
<td>48</td>
</tr>
<tr>
<td>Edmond \textit{et al.}\textsuperscript{43}</td>
<td>1994–1995</td>
<td>Australia</td>
<td>Emergency department</td>
<td>Derivation</td>
<td>Prospective single centre</td>
<td>Throat culture</td>
<td>Pharyngitis on examination by junior medical staff</td>
<td>Unclear</td>
<td>271</td>
<td>21</td>
</tr>
<tr>
<td>McIsaac \textit{et al.}\textsuperscript{40}</td>
<td>1995–1997</td>
<td>Canada</td>
<td>Family medicine centre</td>
<td>Derivation, internal validation</td>
<td>Prospective single centre</td>
<td>Throat culture</td>
<td>New upper respiratory tract infection with a recorded diagnosis of pharyngitis or tonsillitis</td>
<td>3–14*</td>
<td>94</td>
<td>36.2*</td>
</tr>
<tr>
<td>Attia \textit{et al.}\textsuperscript{44}</td>
<td>1999–2000</td>
<td>USA</td>
<td>Emergency department, 2 pediatric outpatient clinics</td>
<td>Update, external validation</td>
<td>Prospective multicentre</td>
<td>Throat culture</td>
<td>Acute pharyngitis</td>
<td>1–18</td>
<td>587</td>
<td>37</td>
</tr>
<tr>
<td>Joachim \textit{et al.}\textsuperscript{37}</td>
<td>2007–2008</td>
<td>Brazil</td>
<td>2 emergency departments, 1 medical unit</td>
<td>Update, external validation</td>
<td>Prospective multicentre</td>
<td>Throat culture or rapid antigen detection testing</td>
<td>Acute pharyngitis</td>
<td>0–15</td>
<td>356</td>
<td>33</td>
</tr>
</tbody>
</table>

\( ^* \text{Studies involved both adults and children; the numbers reported here are for the pediatric populations, as defined by the authors.} \)
<table>
<thead>
<tr>
<th>Clinical prediction rule</th>
<th>Presentation</th>
<th>Clinical predictors</th>
<th>Suggested course of action in the original study</th>
<th>Present validation study</th>
</tr>
</thead>
</table>
| Forsyth36 List           | Temperature, enlarged tender nodes (1 or ≥ 2), exudate (1 patch, ≥ 2 patches or confluent, purulent exudate), infectious mononucleosis syndrome (mucoid exudate and multiple small nodes), influenzal syndrome (fever and myalgia but no exudate or adenitis) | Clinically nonstreptococcal: culture; treat symptomatically  
Clinically “questionable”: culture; treat symptomatically  
Clinically streptococcal: do not culture; give penicillin orally | NA |
| Breese39 Score           | Month in which the patient is seen, age, leukocyte count, fever ≥ 100.5°F, sore throat, cough, headache, abnormal pharynx (redness, swelling, exudate, petechiae, “doughnut lesions” or ulcerations), abnormal cervical nodes (very enlarged without tenderness or if palpable and tender) | No clear course of action suggested  
≤ 25: no rapid antigen detection testing, no antibiotic treatment  
26–31: antibiotic with positive test result  
≥ 32: no testing, antibiotic treatment | |
| Fujikawa41 Score         | Fever > 37.5°C; sore throat, nausea or vomiting, anorexia, absence of cough or rhinorrhea, pharynx (dark red or petechiae), cervical lymph node swelling, tonsillar exudate (white spotted), rash (scarlet fever-like, erythema or urticaria), strawberry tongue or marked papillae | No clear course of action suggested | NA |
| Wald42 Score             | Age, season, fever ≥ 38.3°C, adenopathy (cervical lymph nodes ≥ 1 cm or tender to palpation), pharyngitis (erythema, swelling or exudate of pharynx or tonsils), no upper respiratory symptoms (rhinorrhea, cough or conjunctivitis) | No clear course of action suggested  
≤ 1: no rapid antigen detection testing, no antibiotic treatment  
2–4: antibiotic with positive test result  
≥ 5: no testing, antibiotic treatment | |
| Edmond38 Decision tree   | Age, scarlatiniform rash, pharyngotonsillitis (diffuse pharyngeal erythema and swollen, edematous tonsils), tender cervical nodes | Risk < 20%: consider symptomatic treatment only  
Risk 20%–60%: culture with or without penicillin therapy  
Risk > 60%: penicillin therapy | > 60%: no testing, antibiotic treatment |
| McIsaac40 Score          | Temperature > 38°C, no cough, tender anterior cervical adenopathy, tonsillar swelling or exudate, age | 0–1: no culture or antibiotic required  
2–3: culture all; treat only if result is positive  
≥ 4: culture all or treat with penicillin on clinical grounds | 0–1: no rapid antigen detection testing, no antibiotic treatment  
2–3: antibiotic with positive test result  
≥ 4: no testing, antibiotic treatment |
| Attia43 Score            | Scarlatiniform rash, moderate to severe tonsillar swelling, moderate to severe tenderness and enlargement of cervical lymph nodes, absence of moderate to severe coryza | No clear course of action suggested  
0: no rapid antigen detection testing, no antibiotic treatment  
1–3: antibiotic with positive test result  
≥ 4: no testing, antibiotic treatment | |
| Joachim37 Score          | Age, bacterial signs (tender cervical node, headache, petechiae on the palate, abdominal pain, sudden onset (< 12 h)), viral signs (conjunctivitis, coryza, diarrhea) | ≤ 2: no rapid antigen detection testing, symptomatic treatment  
3: antibiotic with positive result on rapid antigen detection testing  
≥ 4: no rapid antigen detection testing, antibiotic treatment | ≥ 4: no testing, antibiotic treatment |

Note: NA = not applicable (the clinical prediction rule could not be validated).
prediction rule of Attia and colleagues), but this group comprised only 6% of patients in our cohort.

**Interpretation**

Several clinical prediction rules are available for diagnosing childhood pharyngitis. Most of these rules are clinical scores based on the same panel of clinical findings classically associated with group A streptococcal or viral cause. On external validation, none of the rules-based selective testing strategies showed sufficient accuracy, and none were able to identify patients at low or high risk whose condition could be managed without microbiologic testing.

Our primary aim was to identify a rules-based selective testing strategy that would have a sensitivity and specificity of more than 85%. For the definition of such a target zone of accuracy, one needs to account for both end points in the diagnosis of group A streptococcal pharyngitis. The first is diagnostic sensitivity. Patients with false-negative test results might show complications of streptococcal pharyngitis. However, evidence supporting antibiotic treatment to prevent such complications in high-income settings is limited. From a Cochrane review, the number needed to treat with antibiotics to prevent 1 case of quinsy is about 50, but a recent study in the United Kingdom calculated the number needed to treat at about 4000. Acute rheumatic fever seems to have almost disappeared in high-income settings, and all clinical trials done after 1975 (6 trials including 2484 patients) did not show any efficacy of antibiotics in reducing its incidence, with no case of acute rheumatic fever in either arm.

The second end point in the diagnosis of streptococcal pharyngitis is diagnostic specificity. Patients with false-positive test results might receive unnecessary antibiotics, thus contributing to the antimicrobial resistance. In outpatient settings in the United States, respiratory tract infections account for more than 75% of antibiotic use in children under 12 years of age, and pharyngitis is the most frequent diagnosis associated with prescription of antibiotics to children aged 6–12 years. Because of the low impact of antibiotic treatment of streptococcal pharyngitis in children in most industrialized countries and the emergence of antibiotic-resistant bacteria, some northern European countries (e.g., The Netherlands) do not recommend antibiotic treatment for children with pharyngitis, or only for the small number of children with a severe infection or an increased risk of complications.

We considered a rules-based selective strategy clinically relevant only if it allowed for rapid antigen detection tests to be used in less than 80% of cases. In our validation study, the selective testing strategy that had the highest accuracy would have resulted in testing for about 85% of children. In addition, the American Academy of Pediatrics recently recommended rapid antigen detection testing for children with a McIsaac score of 2 or more, but this implies that testing should be done in about 90% of cases. If clinicians accept such high rates of testing, they might agree to have all children presenting with pharyngitis undergo testing. Such a policy is currently recommended by European guidelines and was shown to result in a substantial reduction in antibiotic prescription rates.

### Table 3: External validation of clinical prediction rules in terms of diagnostic accuracy and use of rapid antigen detection testing ($n = 676$)

<table>
<thead>
<tr>
<th>Clinical prediction rule</th>
<th>Area under the ROC curve of the score (95% CI)</th>
<th>Diagnostic accuracy of the rules-based strategy</th>
<th>Rapid antigen detection testing*</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breeze39</td>
<td>0.60 (0.55–0.64)</td>
<td>88 (84–92) 82 (78–86) 0.85 (0.82–0.88)</td>
<td>76 (73–80) 0.01</td>
<td></td>
</tr>
<tr>
<td>Wald42</td>
<td>0.58 (0.53–0.62)</td>
<td>94 (91–97) 63 (58–68) 0.78 (0.76–0.81)</td>
<td>65 (61–69) &lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Edmond38</td>
<td>NA</td>
<td>66 (61–72) 86 (82–89) 0.76 (0.73–0.79)</td>
<td>48 (44–52) &lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>McIsaac40</td>
<td>0.56 (0.52–0.60)</td>
<td>94 (92–97) 54 (49–59) 0.74 (0.71–0.77)</td>
<td>52 (48–56) &lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Attia43</td>
<td>0.62 (0.58–0.66)</td>
<td>87 (83–91) 88 (85–91) 0.87 (0.85–0.90)</td>
<td>86 (84–89) 0.99</td>
<td></td>
</tr>
<tr>
<td>Joachim37</td>
<td>0.59 (0.55–0.63)</td>
<td>88 (84–92) 40 (35–45) 0.64 (0.61–0.67)</td>
<td>24 (21–27) &lt; 0.001</td>
<td></td>
</tr>
</tbody>
</table>

Note: CI = confidence interval, NA = not available.

*Use of rapid antigen detection testing corresponds to the number of patients that would undergo testing after the application of the clinical prediction rule.
†One-sided exact binomial probability test of whether absolute use of rapid antigen detection testing was < 80%.
In our validation study, clinical prediction rules were unable to identify patients at low or high risk in whom testing for group A streptococcus could be avoided. Our results are consistent with those of previous validation studies (Appendix 12, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140772/-/DC1). A recent large-scale validation of the McIsaac score involving 64,789 children aged 3–14 years in the US showed probabilities of 17% in children with a score of 1 and 55% in children with a score 4 or more, which might be far from the misclassification rates clinicians would consider acceptable. Differences in probability by risk level across external validation studies could be explained by differences in study settings and participant characteristics, as well as the low reproducibility of signs and symptoms of streptococcal pharyngitis (i.e., interobserver agreement).

Limitations
About 20% of the children from our validation cohort had missing data for at least 1 clinical predictor. However, we used multiple imputations, which resulted in less bias than the complete-case approach. In a sensitivity analysis, the secondary exclusion of patients with missing data produced slightly higher estimates of diagnostic accuracy than multiple imputations (Appendix 5).

We validated clinical prediction rules using the rapid antigen detection test as a stand-alone test without throat culture, which is the current recommendation of the European Society for Clinical Microbiology and Infectious Diseases.

Several studies have found this strategy to be the most cost-effective. However, whether such tests are sufficiently accurate to replace throat culture in clinical practice remains a matter of debate.

All steps for undertaking the systematic review were done by 2 reviewers, but not independently, increasing the risk that some clinical prediction rules were missed or that mistakes occurred when extracting the data. A further limitation lies in the absence of an a priori sample size calculation. One of the clinical prediction rules met our target zone of accuracy based on the point estimates alone (Attia’s rule), but it was considered insufficient because the boundaries of the confidence intervals for sensitivity and specificity went across the prespecified limits for significance. This could be due to lack of power, and our results should be considered with caution until they are confirmed with a larger sample of patients.

Finally, although this prospective study was originally designed for the purpose of validating existing clinical prediction rules, clinical data were collected before performing the systematic review of the literature. The collection of the clinical variables required for validation was not customized to evaluate the full range of rules we finally included. Some clinical prediction rules had to be adapted because of discrepancies between our clinical variables and the specific definitions originally provided by the authors.

Conclusion
The clinical relevance of clinical prediction rules for triaging children who should undergo group A streptococcal testing was poor in this study, in which both high sensitivity and specificity were expected. Further prospective validation studies comparing multiple clinical prediction rules head-to-head are needed to confirm our results. Future efforts should also aim to better define misclassification rates and the threshold for rapid antigen detection testing that clinicians would consider acceptable.

References
27. Adequate, and measuring and reducing errors.


Affiliations: Obstetrical, Perinatal and Pediatric Epidemiology Research Team (Cohen J.F., Chalumeau), Research Center for Epidemiology and Biostatistics Sorbonne Paris Cité, Paris Descartes University, Paris, France; Department of Pediatrics (Cohen J.F., Chalumeau), Necker-Enfants-Malades Hospital, Assistance Publique-Hôpitaux de Paris, Paris Descartes University, Paris, France; Association Clinique et Thérapeutique Infantile du Val-de-Marne (Cohen R., Levy, Benani), Saint-Maur-des-Fossés, France; Department of Microbiology (Cohen R.), Centre Hospitalier Intercommunal de Créteil, Créteil, France; Clinical Research Center (Levy), Centre Hospitalier Intercommunal de Créteil, Créteil, France; Association Française de Pédiatrie Ambulatoire (Thollot), Essey-lès-Nancy, France; Department of Microbiology (Bidet), Robert Debré Hospital, Assistance Publique-Hôpitaux de Paris, Paris Diderot University, Sorbonne Paris Cité, Paris, France

Contributors: Jérémie Cohen and Martin Chalumeau had full access to all data and take responsibility for the integrity of the data and the accuracy of the analyses. Jérémie Cohen, Robert Cohen, Corinne Levy, Philippe Bidet and Martin Chalumeau conceived and designed the study. Robert Cohen, Franck Thollot, Mohammed Benani and Philippe Bidet acquired the data. Jérémie Cohen and Martin Chalumeau performed the statistical analysis. Jérémie Cohen wrote the first draft of the manuscript, which was revised for important intellectual content by all of the authors. All of the authors read and approved the final version submitted for publication and agree to act as guarantors of the work.

Funding: Jérémie Cohen is a doctoral student supported by educational grants from Agence Régionale de Santé d’Ile-de-France, Laboratoires Guigoz - Société Française de Pédiatrie – Groupe de Pédiatrie Générale – Groupe de Recherches Epidémiologiques en Pédiatrie and research grants from the French Ministry of Health (Régional 2012, 12089) and Assistance Publique-Hôpitaux de Paris (Année Médaille de l’Internat 2013). The clinical study for external validation was internally funded by the Association Clinique et Thérapeutique Infantile du Val-de-Marne and externally by Dectrapharm, manufacturer of the rapid antigen detection test. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Acknowledgments: This article is dedicated to the memory of Professor Edouard Bingen (1946–2012), Department of Microbiology, Robert Debré Hospital and Paris Diderot University. The authors thank A. Liboz, Dr. M. Boucherat, Dr. F. de La Rocque, I. Ramay, D. Menguy and M. Fernandes, and the physician investigators who participated in the study: Dr. F. Corrard, Dr. P. Deberdt, Dr. A. Elbez, Dr. M. Goldrey, Dr. J. Gosselin, Dr. M. Koskas, Dr. P. Martin, Dr. A.S. Michot, Dr. N. D’Ovidio, Dr. D. Qutob, Dr. C. Romain, Dr. O. Romain, Dr. C. Schlemmer and Dr. A. Wollner. The authors also thank R. Spijker for his assistance on the literature search.

Data sharing: No additional data are available.