Large-Scale Validation of the Centor and McIsaac Scores to Predict Group A Streptococcal Pharyngitis

Andrew M. Fine, MD, MPH; Victor Nizet, MD; Kenneth D. Mandl, MD

Background: The Centor and McIsaac scores guide testing and treatment for group A streptococcal (GAS) pharyngitis in patients presenting with a sore throat, but they were derived on relatively small samples. We perform a national-scale validation of the prediction models on a large, geographically diverse population.

Methods: We analyzed data collected from 206,870 patients 3 years or older who presented with a painful throat to a United States national retail health chain from September 1, 2006, to December 1, 2008. Main outcome measures were the proportions of patients testing positive for GAS pharyngitis according to the Centor and McIsaac scores (both scales, 0-4).

Results: For patients 15 years or older, 23% (95% CI, 22%-23%) tested positive for GAS, including 7% (95% CI, 7%-8%) of those with a Centor score of 0; 12% (95% CI, 11%-12%) of those with a Centor score of 1; 21% (95% CI, 21%-22%) of those with a Centor score of 2; 38% (95% CI, 38%-39%) of those with a Centor score of 3; and 57% (95% CI, 56%-58%) of those with a Centor score of 4. For patients 3 years or older, 27% (95% CI, 27%-27%) tested positive for GAS, including 8% (95% CI, 8%-9%) of those testing positive with a McIsaac score of 0; 14% (95% CI, 13%-14%) of those with a McIsaac score of 1; 23% (95% CI, 23%-23%) of those with a McIsaac score of 2; 37% (95% CI, 37%-37%) of those with a McIsaac score of 3; and 55% (95% CI, 55%-56%) of those with a McIsaac score of 4. The 95% CIs overlapped between our retail health chain–derived probabilities and the prior reports.

Conclusion: Our study validates the Centor and McIsaac scores and more precisely classifies risk of GAS infection among patients presenting with a painful throat to a retail health chain.

Published online May 7, 2012.

See Invited Commentary at end of article

The McIsaac score,6,7 derived from 521 patients from a university-affiliated family practice in Toronto, Ontario, Canada, and validated on 621 patients from 49 Ontario communities, adjusts the Centor...
the Institute for Clinical Systems Improvement. According to Centor and McIsaac scores, and test results. The Children’s Hos-
data set included only patient visits where there was complete
and treat only those patients with a positive test for GAS. The
matory testing used for patients whose rapid test is negative),
antigen testing on all patients with pharyngitis (with confir-
manship on algorithm-driven care. The clinicians enter these codified
data in real time, and the information is stored in a common
database across all clinic locations. MinuteClinic providers have
demonstrated greater than 99% adherence to an established acute
pharyngitis protocol, the “Strep Pharyngitis Algorithm,” from
the Institute for Clinical Systems Improvement. According to
this algorithm, medical providers collect structured informa-
tion about patients’ relevant signs and symptoms, obtain rapid
antigen testing on all patients with pharyngitis (with confir-
matory testing used for patients whose rapid test is negative),
and treat only those patients with a positive test for GAS. The
data set included only patient visits where there was complete
information about age, all signs and symptoms included in the
Centor and McIsaac scores, and test results. The Children’s Hos-

### Table 1. ACP/CDC Guidelines for the Management of Pharyngitis

<table>
<thead>
<tr>
<th>Centor Score</th>
<th>ACP/CDC Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Do not test, do not treat</td>
</tr>
<tr>
<td>1</td>
<td>Do not test, do not treat</td>
</tr>
<tr>
<td>2</td>
<td>Treat if rapid test result positive</td>
</tr>
<tr>
<td>3</td>
<td>Option 1: treat if rapid test result positive</td>
</tr>
<tr>
<td>4</td>
<td>Treat empirically</td>
</tr>
</tbody>
</table>

Abbreviations: ACP, American College of Physicians; CDC, Centers for Disease Control and Prevention.

The CDC advocates the ACP guideline based on the Centor score for management of acute pharyngitis in adults. To calculate the Centor score, patients receive 1 point for each of the following: fever; absence of cough; presence of tonsillar exudates; and swollen, tender anterior cervical nodes. Based on these signs and symptoms, the Centor score is calculated (0-4). The McIsaac score adjusts the Centor score to account for the increased incidence of group A Streptococcus infection in children and decreased incidence in older adults by adding 1 point to the Centor score for those younger than 15 years and subtracting 1 point for those 45 years or older.

score for the patient’s age. Since younger patients are more likely to have GAS than older patients, the McIsaac score is calculated by adding 1 point to the Centor score for patients aged 3 to 14 years and subtracting 1 point for those 45 years or older. Because clinical prediction models may perform poorly when applied to new settings, it is important to validate them on different populations and over time. Furthermore, despite endorsement from CDC and ACP-ASIM, the clinical scores have gained poor traction in clinical practice, perhaps in part owing to the perception that the scores were derived from a relatively small sample. Herein, we analyze a geographically diverse population of patients who presented with sore throat to MinuteClinic, a large retail health chain, to perform the largest validation studies of the Centor and McIsaac scores.

# METHODS

## STUDY DESIGN

We analyzed retrospective data collected from patients tested for GAS pharyngitis when they presented with a painful throat from September 1, 2006, to December 1, 2008, to MinuteClinic, a large, national retail health chain with over 500 sites in 26 states. From the retail clinic’s 581 sites, the data set included 238,596 patient encounters across 25 states. In this setting, physician assistants or nurse practitioners collect standardized historical and physical examination information based on algorithm-driven care. The clinicians enter these codified data in real time, and the information is stored in a common database across all clinic locations. MinuteClinic providers have demonstrated greater than 99% adherence to an established acute pharyngitis protocol, the “Strep Pharyngitis Algorithm,” from the Institute for Clinical Systems Improvement. According to this algorithm, medical providers collect structured information about patients’ relevant signs and symptoms, obtain rapid antigen testing on all patients with pharyngitis (with confirmatory testing used for patients whose rapid test is negative), and treat only those patients with a positive test for GAS. The data set included only patient visits where there was complete information about age, all signs and symptoms included in the Centor and McIsaac scores, and test results. The Children’s Hos-

All MinuteClinic locations used the Clinical Laboratory Improvement Amendments–waived QuickVue In-line Strep A test (Quidel Corp). The confirmatory test was a streptococcal DNA probe (74%) or throat culture (26%). Patients were categorized as GAS positive if the finding of either test (rapid or confirmatory) was positive.

## STATISTICAL ANALYSIS

Predictor variables and covariates were developed for age, sex, history of fever in previous 24 hours, history of exposure to someone with GAS pharyngitis, presence of cough, duration of pharyngitis symptoms (days), presence of erythematous tonsils, presence of tonsillar exudates, presence of swollen tonsils, presence of swollen posterior cervical lymph nodes, presence of swollen anterior cervical lymph nodes, and absence of cough. The McIsaac score was calculated for all patients 3 years or older by adding 1 point to the Centor score for those younger than 15 years and by subtracting 1 point from the Centor score for those 45 years or older.

Two approaches were taken to validate the scores. First, we compared the likelihood of GAS pharyngitis by clinical score in the MinuteClinic patients to the likelihood of GAS pharyngitis by clinical score in the published data. Second, we applied logistic regression to the MinuteClinic data to derive new prediction models, maintaining the same parameter that they be limited to no more than 4 clinical variables. The 4 chosen variables derived from the cohort of patients 15 years or older were then compared with the 4 variables that compose the Centor score.

## CALCULATION OF GAS PROBABILITIES

The percentage of patients 15 years or older in the retail health data who tested positive for GAS by Centor score (0-4) was calculated and compared with the original report by Centor et al. and with the validation study by Wigton et al. The percentage of patients 3 years or older in the retail health data who tested positive for GAS by McIsaac score (0-4) was calculated and compared with the McIsaac studies. Ninety-five percent
CIs were calculated for the proportion of patients testing positive at each score. The 95% CIs around the proportion testing positive by score in the retail health data were compared with the 95% CIs in the Centor\textsuperscript{4,18} and McIsaac\textsuperscript{6,7} studies.

**SELECTION OF VARIABLES**

Variables included in the Centor and McIsaac scores as well as variables not included in the scores were examined to determine the best predictors of GAS pharyngitis among the MinuteClinic patients. Univariate and multivariate analyses were performed to identify predictors of GAS pharyngitis. Significance of association of categorical variables with GAS pharyngitis was tested by the χ\textsuperscript{2} test. In the multivariate analysis, candidate predictors were entered into a stepwise logistic regression to identify independent predictors of patients with GAS pharyngitis. P value cutoffs for entry and departure for the multivariate regression models were .25 and .10, respectively. For the purpose of simplicity and usability and to facilitate comparison with the prior studies, the final model was limited to 4 predictor variables and assessed by area under the receiver operator characteristic curve (AUC). Statistical analyses were performed using JMP Pro software, version 9.0.2 (SAS Institute).

**RESULTS**

Of 238,656 patient visits, 5653 were excluded owing to treatment for GAS within the prior month, and an additional 1399 were excluded because the patient age was younger than 3 years, leaving 231,604 patient visits. For patients with multiple visits, only the first visit was included, leaving 206,870 patients to validate the McIsaac score. Of these, 64,789 visits occurred for patients younger than 15 years (31%), leaving 142,081 visits for the validation of the Centor score (Figure). Among the 142,081 retail health visits for patients 15 years or older, 23% (95% CI, 22%-23%) tested positive for GAS, compared with 17% (95% CI, 14%-23%) in the original Centor et al\textsuperscript{4} report. As in the original Centor et al\textsuperscript{4} report. Presence of tonsillar exudates conferred the highest odds of having streptococcal infection (3.1 [95% CI, 3.0-3.2]) followed by swollen anterior cervical lymph nodes (2.2 [95% CI, 2.1-2.3]), history of fever (1.7, [95% CI, 1.7-1.8]) and absence of cough (1.6, [95% CI, 1.5-1.6]).

The overall performance of the model as applied to the retail health data was evaluated by comparing the AUCs. For patients 15 years or older, applying the Centor score to the retail health data yielded an AUC of 0.72. For patients 3 years or older, applying the McIsaac score to the retail health data achieved an AUC of 0.71.

**COMMENT**

We evaluated 2 commonly used prediction models to classify risk of GAS pharyngitis among patients presenting with a painful throat. The purpose of a clinical prediction model is to provide clinicians with a practical and applicable tool to improve medical decision-making, the health of individual patients, and the public health. The Centor score is 1 model that is particularly robust; it has withstood 30 years of changes in diagnostic testing, information technology, and population dynamics.\textsuperscript{10} Our study validated the Centor score in a clinical setting (retail health chain) with a less acutely ill population than is seen in the emergency department setting from which the score was derived. While the Centor score was derived from a relatively small number of patients (n = 286) seen in 1 setting during a single 2-month period, we analyzed data from multiple locations spanning more than 1 calendar year, mitigating the potential impact of sea-

![Patient flow diagram](figure.png)
Table 2. Characteristics of the 206 870 Retail Health Patients With Pharyngitis

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Age $\geq$ 3 Years (n = 206 870)</th>
<th>Age $\geq$ 15 Years (n = 142 081)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAS Positive (n = 56 013)</td>
<td>GAS Negativeb (n = 150 857)</td>
</tr>
<tr>
<td>Age, mean (median) [IQR], y</td>
<td>23 (20) [9-35]</td>
<td>28 (26) [14-39]</td>
</tr>
<tr>
<td>Male sex</td>
<td>22 768 (41)</td>
<td>54 540 (36)</td>
</tr>
<tr>
<td>Fever</td>
<td>30 710 (55)</td>
<td>52 006 (34)</td>
</tr>
<tr>
<td>Absence of cough</td>
<td>40 538 (72)</td>
<td>93 255 (62)</td>
</tr>
<tr>
<td>Anterior cervical lymphadenopathy</td>
<td>42 662 (76)</td>
<td>83 249 (55)</td>
</tr>
<tr>
<td>Tonsillar exudate</td>
<td>21 963 (39)</td>
<td>24 513 (16)</td>
</tr>
<tr>
<td>Tonsillar swelling</td>
<td>34 525 (62)</td>
<td>54 528 (36)</td>
</tr>
<tr>
<td>Temperature $\leq$ 101°F</td>
<td>3455 (6)</td>
<td>4213 (3)</td>
</tr>
<tr>
<td>Exposure to GAS</td>
<td>19 718 (35)</td>
<td>38 429 (25)</td>
</tr>
<tr>
<td>Lack of rhinorrhea</td>
<td>44 473 (79)</td>
<td>110 666 (73)</td>
</tr>
<tr>
<td>Posterior cervical</td>
<td>5044 (9)</td>
<td>8651 (6)</td>
</tr>
<tr>
<td>Lymphadenopathy</td>
<td>5044 (9)</td>
<td>8651 (6)</td>
</tr>
<tr>
<td>Symptom duration</td>
<td>10 557 (20)</td>
<td>23 449 (17)</td>
</tr>
<tr>
<td>$&lt;24$ h</td>
<td>25 928 (46)</td>
<td>56 172 (38)</td>
</tr>
<tr>
<td>1-2 d</td>
<td>14 436 (26)</td>
<td>43 138 (29)</td>
</tr>
<tr>
<td>3-4 d</td>
<td>5092 (9)</td>
<td>23 563 (16)</td>
</tr>
<tr>
<td>$\geq5$ d</td>
<td>10 557 (20)</td>
<td>23 449 (17)</td>
</tr>
</tbody>
</table>

Abbreviations: GAS, group A Streptococcus; IQR, interquartile range.

a Unless otherwise indicated, data are presented as number (percentage) of patients.

b P < .001 for every entry in this column (positive vs negative).

d The percent of patients testing positive according to clinical scores fell within the 95% CIs of the original study by Centor et al, which was based on fewer than 300 patients. As our research group has shown previously, the recent local incidence of GAS pharyngitis further improves the accuracy of estimating an individual patient’s risk of GAS pharyngitis. Troughs and peaks of GAS pharyngitis outbreaks will occur naturally throughout the year, so the retail health data in our analyses collected over more than 1 year average over those variations and should provide more reliable characterization of the score than did the original study by Centor et al, which was conducted over only 2 months. The AUC is a metric widely used to reflect the overall accuracy of a diagnostic test or overall performance of a clinical prediction model. The AUC of the Centor score in the present retail health population (0.72) was lower than that found in the original 1981 study by Centor et al (0.78) but the same as that found in the validation study by Wigton et al (0.72), arguing for the discriminating validity of this score. Clinical prediction models tend to perform less well in validation studies, but our data are consistent with the model’s performance in other validation studies. While McIsaac et al did not report an AUC with their original data, the McIsaac score performed similarly to the others in the present large data set. The observed proportion of MinuteClinic patients testing positive according to clinical scores fell within the 95% CIs of the Wigton et al and McIsaac et al validation studies (except for McIsaac score 0), sup-
porting the calibration validity of the Centor and McIsaac scores.

Leveraging codified data from retail health clinics where uniform, algorithm-driven care is provided and data are captured in a single electronic medical record, our study demonstrates the strengths of the Centor and McIsaac scores as useful tools in clinical decision making. Though many clinicians in the primary care or emergency medicine setting do not routinely test adult patients who are either very likely or very unlikely to have GAS pharyngitis (ie, those with Centor scores of 0, 1, and 4), because MinuteClinic protocol mandates testing for all patients presenting with a painful throat, a further unique strength of our large validation study is ascertainment of GAS status on all subjects.

Though all clinical and laboratory data were collected prospectively, the analyses were conducted retrospectively. There may be some variability in clinical interpretation of the Centor criteria by the nurse practitioners in the MinuteClinic setting; whether anterior cervical nodes are enlarged, for example, might be more subjective than other criteria such as temperature above 101°F (38.3°C).22 Furthermore, data are not available for calculating interobserver or intraobserver reliability.

Though very useful for diagnosing the presence of GAS, retail health data would be unlikely to detect group C Streptococcus and most other bacterial causes of pharyngitis, including Fusobacterium necrophorum, which may cause severe disease especially in adolescents and young adults.23

All patients in the data set were symptomatic with sore throat, so our analyses do not address the important issue of the asymptomatic streptococcal carrier state. Serologic testing was not performed, so symptomatic patients with a positive GAS test finding were assumed to be true positives, not carriers.

Because these data were collected recently, we could not quantify potential changes in antibiotic uses attributable to the 2002 Infectious Diseases Society of America guideline24 and to the 2001 American College of Physicians guideline.25

Using national-scale and uniform data electronically captured from a retail clinic chain, we have validated the Centor and McIsaac scores as useful and valid tools for diagnosis and treatment of patients with acute pharyngitis.

Accepted for Publication: February 17, 2012. Published Online: May 7, 2012. doi:10.1001/archinternmed.2012.950

Correspondence: Andrew M. Fine, MD, MPH, Division of Emergency Medicine—Main 1, Children's Hospital Boston, 300 Longwood Ave, Boston, MA 02115 (andrew.fine@childrens.harvard.edu).


Financial Disclosure: None reported.

Funding/Support: This study was funded by CDC Mentored Public Health Research Scientist Development Award K01HK000055 (Dr Fine); National Library of Medicine, National Institutes of Health grants R01 LM007677 and G08LM009778 (Dr Mandl); and CDC Public Health Informatics Center of Excellence Award P01HK000088 (Dr Mandl).

Additional Contributions: CVS/Caremark and MinuteClinic provided the data used in this study. However, they had no role in the study design, data analysis, data interpretation, manuscript drafting or revision, or the decision to submit for publication.

REFERENCES

Adolescent and Adult Pharyngitis

More Than “Strep Throat”

In this issue of the Archives, Fine and colleagues report a major validation of clinical prediction rules for predicting group A streptococcal (GAS) pharyngitis. They examine the prediction model that I and my research team first reported in 1981 and the modification (incorporating age into the decision rule) that McIsaac et al reported in 1998. This validation confirms a recent meta-analysis that arrived at the same finding. These models provide a probability of a positive group A β-hemolytic streptococcal culture based on a prevalence estimate and 4-point scoring system. The 4-point system appears to work well for preadolescent pharyngitis, where ultimately the clinician must make a dichotomous decision—GAS infection or a viral infection.

However, as patients enter adolescence and continue growing into young adulthood, the model becomes more controversial; indeed, the 2 US guidelines differ in their approach to adult pharyngitis. To understand why there is less certainty in the diagnosis of pharyngitis in adolescents and older persons, we note several important differences between this population and young children:

1. Adolescents and young adults respond to penicillin treatment of GAS with a 2-day decrease in symptoms, while preadolescents do not appear to show this effect.

2. Group C (and other non–group A) β-hemolytic streptococcal pharyngitis occurs more frequently in adolescents and young adults than in preadolescents. In adolescents and young adults, treatment of group C streptococcal (GCS) pharyngitis results in a 1-day shorter duration of symptoms; and

3. While both age groups develop Epstein-Barr infections, only the adolescents and young adults develop the infectious mononucleosis syndrome; and

4. Fusobacterium necrophorum pharyngitis occurs much more frequently in adolescents and young adults, as does the Lemierre syndrome (a syndrome of internal jugular thrombophlebitis and metastatic infections usually caused by F necrophorum).

Given these differences, should we empirically treat adolescents and young adults with antibiotics if they have a Centor score of 3 or 4? To answer this question we should make explicit the potential benefits and risk of antibiotic therapy.

If a pharyngitis has a bacterial cause (at least with GAS and GCS pharyngitis), then treatment with appropriate antibiotics will decrease symptom duration for adolescents and young adults. We cannot be certain that antibiotics can decrease the duration of fusobacterial pharyngitis. When the patient has bacterial pharyngitis, antibiotics decrease bacterial spread to others.

Treatment of GAS pharyngitis decreases suppurrative complications. Treatment without a GAS diagnosis decreases peritonsillar abscesses; therefore, we can deduce that treating fusobacterial pharyngitis should decrease peritonsillar abscesses. Recent studies suggest that F necrophorum is the most common anaerobe causing peritonsillar abscesses. While we cannot prove that treating fusobacterial pharyngitis will decrease the incidence of Lemierre syndrome, that conclusion seems likely. Treating GAS pharyngitis decreases the incidence of the nonsuppurative complication acute rheumatic fever. It may also decrease the risk of acute glomerulonephritis. Group C Streptococcus clearly can cause glomerulonephritis and may even cause some cases of acute rheumatic fever. We are not aware of any nonsuppurative complications from fusobacterial pharyngitis.