Published in final edited form as:
*Ann Intern Med*. 2011 September 20; 155(6): . doi:10.7326/0003-4819-155-6-201109200-00002.

# Improved Diagnostic Accuracy of Group A Streptococcal Pharyngitis Using Real-Time Biosurveillance

**Andrew M. Fine, MD, MPH**[1,2], **Victor Nizet, MD**[3], and **Kenneth D. Mandl, MD, MPH**[1,2,4,5]

[1]Division of Emergency Medicine, Department of Medicine, Children's Hospital Boston, MA

[2]Department of Pediatrics, Harvard Medical School, Boston MA

[3]Department of Pediatrics and Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California

[4]Children's Hospital Informatics Program, Harvard-MIT Health Sciences and Technology, Boston, MA

[5]Center for Biomedical Informatics, Harvard Medical School, Boston, MA

## Abstract

**Background—**Clinical prediction rules do not incorporate real time incidence data to adjust estimates of disease risk in symptomatic patients.

**Objective—**To measure the value of integrating local incidence data into a clinical decision rule for diagnosing group A streptococcal (GAS) pharyngitis in patients age 15 years and older.

**Design—**Retrospective analysis of clinical and biosurveillance predictors of GAS pharyngitis.

**Setting—**Large U.S.-based retail-health chain.

**Patients—**82,062 patient visits for pharyngitis.

**Measurements—**Accuracy of the Centor score, was compared with that of a biosurveillance-responsive score, essentially an adjusted Centor score based on real-time GAS pharyngitis information from the 14 days prior to a patient's visit – the recent local proportion positive (RLPP).

**Results—**Increased RLPP correlated with likelihood of GAS pharyngitis ($r^2$ =0.79, p<0.001). Local incidence data enhanced diagnostic models. For example, when RLPP >0.30, managing patients with Centor scores of 1 as if scores were 2 would identify 62,537 previously missed patients annually while misclassifying 18,446 patients without GAS pharyngitis. Decreasing the score of patients with Centor values of 3 by one point for RLPP <0.20, would spare unnecessary antibiotics for 166,616 patients while missing 18,812 true positives.

**Limitations—**Analyses were conducted retrospectively. Real time regional GAS pharyngitis data are generally not yet available to clinicians.

**Conclusions—**Incorporating live biosurveillance data into clinical guidelines for GAS pharyngitis and other communicable diseases should be considered to reduce missed cases when

Correspondence to Andrew M. Fine, MD, MPH, Division of Emergency Medicine – Main 1, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA, 02115. Phone 617-355-9696. Fax 617-730-0335. andrew.fine@childrens.harvard.edu.
**Mailing addresses for other authors:** Kenneth D. Mandl, MD, MPH. Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115; Victor Nizet, MD, Division of Pediatric Pharmacology and Drug Discovery, UCSD School of Medicine, Cellular & Molecular Medicine East, Room 1066, 9500 Gilman Drive Mail Code 0687, La Jolla, CA, 92093-5611.

**Conflict of interest:** All authors have declared that no competing interests exist.

the contemporaneous incidence is elevated and spare unnecessary antibiotics when the contemporaneous incidence is low. Delivering epidemiologic data to the point of care will enable the use of real-time pre-test probabilities in medical decision-making.

**Primary Funding Source—**The Mentored Public Health Research Scientist Development Award K01 HK000055 from the Centers for Disease Control and Prevention and R01 LM007677 from the National Library of Medicine, National Institutes of Health.

## Introduction

For communicable diseases, the risk of infection depends on local incidence (1, 2). However, clinicians rarely can access those data at the point of care and do not use formal quantitative approaches to adjust estimates of disease probability based on local incidence (3). The maturation of real-time infectious disease surveillance systems (3–7), coupled with the increased adoption of electronic health records, offers opportunities to improve clinical decision-making by incorporating up-to-date information about local disease incidence into decision support algorithms. Even for diseases like influenza where robust surveillance exists and results are widely publicized, there are no quantitative, automated methods to adjust decision support with real-time incidence data (8). Here we evaluate use of real-time biosurveillance data for management of adults and older adolescents with pharyngitis.

Group A *Streptococcus* (GAS) is the most common bacterial agent causing acute pharyngitis, with 600 million cases occurring annually worldwide (9). The disease occurs across the age spectrum, with a peak incidence in school-age children (5 to 15 years). Timely antibiotic treatment reduces acute rheumatic fever, suppurative complications, symptom duration, missed work and school days, and transmission (10). Because physical examination is an unreliable method to diagnose GAS pharyngitis, the American College of Physicians (ACP) and the Centers for Disease Control and Prevention (CDC) advocate using the Centor score to classify GAS pharyngitis risk in adults (11, 12). Clinicians assign the Centor score (0–4) with one point each for presence of exudates, history of fever, presence of swollen anterior cervical lymph nodes, and absence of cough. Despite evidence-based guidelines patients with pharyngitis are often misclassified, leading to inappropriate antibiotics for those with viral disease and under-treatment of those with *bona fide* GAS pharyngitis. The empiric strategy proposed by the ACP leads, in some cases to unnecessary antibiotics (13, 14), contributing to antimicrobial resistance.

GAS pharyngitis occurs sporadically, with spatial and temporal fluctuations in incidence and occasional outbreaks (15–18), reflecting a dynamic epidemiology of causative strains (19). Even when guidelines are followed, the accuracy of rapid GAS pharyngitis testing is affected by the pre-test probability of disease, which is related to the likelihood of exposure. In temperate climates, GAS pharyngitis may peak in the winter and early spring (16), but clinicians currently lack accurate methods to recognize the onset, duration, or magnitude of local outbreaks from year to year. Here we compare the diagnostic accuracy of the Centor score adjusted with real-time contemporaneous local disease incidence (biosurveillance) data vs. the Centor score alone to identify GAS pharyngitis. We limited the clinical analysis to patients age 15 years and older, for whom the Centor score and ACP guidelines are intended.

## Methods

### Participants and Setting

We retrospectively analyzed data collected from patients tested for GAS pharyngitis when they presented with pharyngitis to MinuteClinic, a large, national retail health chain with over 500 sites in 26 states (20–23). MinuteClinic provided data for the following time

periods: October 1, 2006 – February 6, 2008 and August 14, 2008 – October 31, 2008. Patients were included if they presented with a primary complaint of pharyngitis and were tested for GAS pharyngitis, or if they had symptoms of pharyngitis and were tested for GAS pharyngitis. MinuteClinic practice is to care only for non-septic appearing patients. Patients could contribute more than one encounter. We did not exclude patients who were pregnant, known to have been treated for streptococcal pharyngitis in the prior month, already on antibiotics or had co-morbid diseases. Nurse practitioners or physician assistants routinely collected a consistent set of historical and physical exam elements stored in a common database across all locations. These providers have demonstrated over 99% adherence to an established acute pharyngitis protocol – the "Strep Pharyngitis Algorithm" from the Institute for Clinical Systems Improvement (24); providers collect structured data on signs and symptoms, obtain rapid GAS pharyngitis tests for all pharyngitis patients (with confirmatory testing for those who test negative) and treat only those testing positive. The data set was limited to patients with complete information for the following variables: visit date, MinuteClinic location, age, signs and symptoms included in the Centor score, and test results.

## Test Methods

All sites used the CLIA (Clinical Laboratory Improvement Amendments)-waived QuickVue InLine Strep A test (Quidel Corporation, San Diego, CA). The confirmatory test was either a throat culture (42%) or strep DNA probe (58%). Patients were considered GAS pharyngitis positive if either the rapid or confirmatory test was positive (25).

## Statistical Analysis

Statistical analysis was restricted to the nine MinuteClinic markets with at least 7,000 patient visits each for pharyngitis over the study period, encompassing 132,821 patient encounters across six states (Georgia, Indiana, Maryland, Minnesota, North Carolina and Tennessee). For clinical analyses, we included visits from patients at least 15 years old (n=82,062) with two-thirds (n=54,981) selected randomly for derivation and the rest (n=27,081) for validation. We included data from patients younger than age 15 years (n=50,759) when calculating the overall local incidence data, because they contribute to the epidemiologic context of the clinically analyzed population.

To enable integration of contemporaneous, local GAS pharyngitis epidemiologic data, we created a biosurveillance variable reflecting disease incidence called *recent local proportion positive* (RLPP), a moving window defined as:

$$\mathrm{RLPP}=\frac{Number of patients testing\ GAS\ pharyngitis\ positive\ in\ market\ A\ in\ previous\ 14\ days}{Total\ number of patients\ tested\ for\ GAS\ pharyngitis\ in\ market\ A\ in\ previous\ 14\ days}$$

To calculate the 14-day RLPP for a patient seen on October 15, 2007, for example, we divided the number of positive GAS pharyngitis tests by the number of GAS pharyngitis tests sent in that market from October 1–14, 2007. We calculated 3, 7 and 14-day RLPPs and compared them using Pearson's correlation coefficients.

Patients were grouped by Centor score (0–4) and RLPP (in intervals of 0.01). For example, we calculated the proportion of patients with a Centor score of 2 who tested GAS pharyngitis positive when the RLPP was 0.30. We performed this calculation for all combinations of Centor scores (0–4) and RLPP (0.10–0.40) when at least 40 patients shared a combination. For each Centor score, we plotted the proportion of patients testing positive as a function of RLPP. Linear regression was used to determine the strength of association

between the RLPP and GAS pharyngitis positivity for each Centor score, and Pearson coefficient was used for correlation.

We considered the ACP guideline recommending those with a score of 1 should neither be tested nor treated for GAS pharyngitis, those with a score of 2 should be tested and treated only if they test positive, and those with a score of 3 should be treated empirically (one of two ACP options). For patients with a Centor score of 1 or 2, we evaluated a biosurveillance-responsive score constructed by adding a point to the score of patients when the RLPP exceeded the following thresholds: 0.20, 0.25, 0.30, and 0.35. When the RLPP exceeded the threshold and triggered an increase in the score, we calculated the number of patients in that group (n) who would be correctly and incorrectly reclassified as GAS pharyngitis positive. The assumption is that testing is 80% sensitive and 95% specific (26). We subtracted the number incorrectly reclassified from the number correctly reclassified to determine the net reclassification, and divided it by the number of patients (n) to determine the net percent reclassification at each RLPP threshold. To facilitate comparison across the different RLPP thresholds, we calculated the number of patients reclassified correctly and incorrectly for hypothetical cohorts of 1000 patients. To determine the number needed to test to identify each additional case of GAS pharyngitis, we divided the additional number of patients tested by the additional number correctly reclassified as GAS pharyngitis positive at each threshold.

We calculated national estimates of the number of patients where management would have been different when using the biosurveillance-responsive approach compared to the traditional Centor score approach. To calculate the national impact on an estimated 10.5 million annual national pharyngitis visits (27), we extrapolated the relevant distributions from the MinuteClinic data for age (62% were 15 years or older), Centor score (i.e. 32% of patients had a Centor score of 1) and RLPP (i.e. 21% of patients presented when the RLPP exceeded 0.30). We examined the effect of reducing the score by one for those with a Centor score of 2 or 3 if the RLPP was below the following thresholds: 0.15, 0.20, and 0.25. We calculated the number of patients and number per 1000 correctly and incorrectly reclassified as GAS pharyngitis negative, the net number and net percent reclassified, and the estimated number of patients nationally whose management would have been altered. The Children's Hospital Boston Committee on Clinical Investigation approved this database analysis. We performed statistical analyses using JMP software, version 8 (SAS Institute, Cary, North Carolina).

## Role of the Funding Source

This work was supported by the Centers for Disease Control and Prevention, and by the National Library of Medicine, National Institutes of Health. Only the authors had a role in the study design, data analysis, data interpretation, writing of the manuscript, and the decision to submit this manuscript for publication.

## Results

Table 1 shows patient characteristics for the derivation set (n=54,981), including distribution by gender, age, clinical findings and Centor score. Two-thirds of patients had Centor scores of 1–2, and 68% of patients were female. Thirteen percent of the patients were 15–18 years old. Most patients (91%) had a single encounter, and seven percent had two encounters. In the validation set, the distribution of patients was the same as in the derivation set for age, gender, presence of fever, presence of swollen anterior cervical nodes, absence of cough, presence of tonsillar exudates, and distribution of Centor score. A median number of 7,289 patients were tested per month (IQ range 4365–8602). With regard to volume over time,

29,826 patients were tested in the first quarter of the study, 32,143 in the second, 36,156 in the third, and 34,696 in the fourth. The number of patient-visits, n, for the entire study population were: Georgia 1 (n=7777), Georgia 2 (n=9797), Maryland (n=9720), North Carolina 1 (n=12,236), Indiana (n=8901), Tennessee (n=15,365), North Carolina 2 (n=10,122), Minnesota 1 (n=30,391), and Minnesota 2 (n=27,972).

The proportion of all patients testing positive varied across time and location, demonstrating no obvious predictable GAS pharyngitis season (Figure 1). For example, during the week of December 24, 2007, the proportion positive in three markets was below 20%, in four was 20–29%, and in two was above 30%. Three, seven, and 14-day RLPPs were strongly correlated (14 vs 7: $r^2$= 0.79, p<0.001, 7 vs 3: $r^2$=0.63, p < 0.001, 14 vs 3, $r^2$= 0.48, p<0.001), so we used 14-day RLPP for subsequent analyses because it provides a realistic time frame to generate reliable, contemporaneous local GAS pharyngitis data.

Overall, 25% of all patients tested GAS pharyngitis positive in the derivation and validation sets, higher than the 17% in the original Centor study, but comparable to Wigton's validation study (26%) (11, 28). For patients with Centor scores 1–4, the proportion testing positive is lowest when the RLPP is low, and increases with rising RLPP (p values for slopes <0.001). Figure 2 illustrates the proportion testing positive plotted by RLPP, grouped by Centor score. Each point on the graph represents a group of patients with an identical Centor score/RLPP dyad. Overall, a patient with a Centor score of 3 is more likely than a patient with a Centor score of 2 to test GAS pharyngitis positive (43% – 25%), but this changes under particular epidemiologic conditions. To illustrate, 350/1053 (33%, 95% CI 30%–36%) patients with a Centor score of 2 test positive when the RLPP is above 0.35, compared to 109/328 (33%, 28%–38%) with a Centor score of 3 when the RLPP is less than 0.15. Figure 3 (see Appendix) shows a similar graphical distribution for the same analysis using the validation set.

We measured the impact of adding one point to a Centor score of 1 when the RLPP exceeded specific thresholds. Hypothetically, testing 1000 patients when the RLPP is > 0.30, for example, would correctly reclassify 139 and incorrectly reclassify 41 patients as positive, a ratio exceeding 3:1 (Table 2).

By extrapolation, we estimated that approximately 449,908 of the 10.5 million pharyngitis visits per year, would occur among those 15 years or older, with a Centor score of 1 when the RLPP exceeded 0.30. The biosurveillance approach to guide management of these patients would identify 62,537 additional GAS pharyngitis cases in the U.S. annually, while treating an additional 18,446 patients without GAS pharyngitis. The number needed to test to detect each additional case of GAS pharyngitis is 7.2. Table 4 (see appendix) displays the corresponding analyses for the validation set: 60,048 additional cases would be identified, while 18,103 patients without GAS pharyngitis would be treated. In the validation set, the number needed to test is 7.4.

We examined outcomes generated by adding one point to the Centor score of adults and older adolescents with a score of 2 (n=18,942) at specific RLPP cutoffs (Table 2). Incrementing the score when the RLPP >0.30 and empirically treating a simulated cohort of 1000 of these patients would correctly reclassify 62 but incorrectly reclassify 657 patients as positive. Extrapolating, this approach correctly identifies 29,450 additional GAS pharyngitis cases in the U.S. annually, but at a cost of inappropriately treating 312,077 patients without GAS pharyngitis.

Next, we tabulated outcomes generated by subtracting one point from the Centor score of all adults and older adolescents with a score of 3 (n=10,056) or 2 (n=18,942) at defined RLPP cutoffs (Table 3). Testing 1000 hypothetical patients with a score of 3 when the RLPP <

0.20, rather than treating them empirically would correctly reclassify 620 but incorrectly reclassify 70 patients as negative. In the U.S., this could spare antibiotics for 166,616 patients (5 million doses) while missing only 18,812 cases annually. Table 5 (appendix) shows the corresponding data from the validation set: 169,637 patients would be spared antibiotics, while 17,632 patients would be missed.

## Discussion

We asked the question, "What if clinicians could integrate real-time incidence data into clinical decision-making for patients?" Retail health data on pharyngitis were ideal to address this question because clinical scores have been validated to estimate GAS pharyngitis risk in adults, the disease is common, there is temporal and spatial variability in disease incidence, and by MinuteClinic protocol, all pharyngitis patients undergo GAS pharyngitis testing, regardless of Centor score. We show that contextualizing the Centor score, using biosurveillance data to calculate the RLPP, improves prediction of GAS pharyngitis positivity in patients presenting with pharyngitis, across all clinical risk categories, and is especially important when the RLPP is at the extremes (high or low). For adults and older adolescents, adjusting management based on a biosurveillance approach could improve health outcomes and efficiency of health care delivery.

The ACP/CDC recommends that adults unlikely to have GAS pharyngitis (Centor 0 or 1) should neither be tested nor treated. We show that adults and older adolescents with a Centor score of 0 are unlikely to have GAS pharyngitis regardless of RLPP; even when the RLPP is elevated, their risk of GAS pharyngitis rarely exceeds 15%. However, our data suggest that during times of elevated RLPP, clinicians should consider testing adults and older adolescents with a Centor score of 1. Adding a point for adults and older adolescents with a Centor score of 1 offers greater overall benefit than adding a point to those with a score of 2. Also, subtracting one point from a score of 3 when the RLPP is low, on balance spares unnecessary empiric antibiotic therapy with a small reduction in case identification.

Practice guidelines traditionally account for clinical features, and to some extent seasonality, without regard for real-time epidemiological disease incidence data. Some clinicians informally incorporate epidemiologic incidence data into their heuristics for medical decision making, but a clinician's knowledge of current disease trends may be prejudiced by recent or unusual cases, leading to cognitive bias and over- or under-estimation of the true incidence of disease (29, 30). The propagation of retail health clinics in geographically diverse areas provided an opportunity to evaluate how local incidence data collected via an unbiased robust, uniform information system can improve clinical decision making.

Developing formal guidelines for incorporating biosurveillance data into the diagnostic and treatment algorithms for GAS pharyngitis, necessitates a discussion of the relative tradeoffs between missed cases and over-treatment of those without true bacterial infection. In one example, increasing the Centor score by one would have resulted in 3.4 additional cases of GAS pharyngitis diagnosed for every additional negative treated. Nevertheless, the public health concern of antibiotic resistance must be weighed against the benefits of preventing complications and morbidities from missed cases.

## Limitations

Patients presenting to retail health clinics may be less acutely ill and have fewer co-morbidities than patients with pharyngitis who present to other types of health care settings, and our findings should not be applied to individuals suspected of having sepsis. Further, though all clinical and laboratory data were collected prospectively, the analyses were conducted retrospectively. There are logistical considerations when integrating the RLPP

into a clinical decision support framework. The RLPP surveillance signal stream would weaken if clinicians tested fewer low-risk patients when the RLPP is low. To preserve the accuracy of the RLPP metric, it would be imperative to maintain a stream of data through the random, intentional testing of a threshold sample of patients weekly, regardless of Centor score and RLPP. This testing approach could provide value in disease management from the population health perspective.

Our data are not able to inform the debate about the importance of the streptococcal carrier state because serologic tests for GAS pharyngitis were not performed. In this analysis, all patients were symptomatic with pharyngitis, and this clinical presentation coupled with positive GAS pharyngitis testing is universally approached as true infection in clinical practice and decision-making algorithms. We did not have information about family or close contacts that tested positive for GAS pharyngitis, but we expect that clinicians would continue to consider this information, if available. Our analyses did not cover all regions of the country, but did include six different states.

While very large sample size is a strength of our study, the dependence on multiple providers' clinical assessments may introduce some variability. However, all information was collected by NPs and PAs trained to evaluate pharyngitis according to a strict, computer-driven protocol. Data were not available to permit calculation of intra- or inter-observer variation in the collection of clinical findings that contribute to the Centor score. We are limited by the absence of data about locations of patients' schools or jobs that might permit more refined epidemiologic modeling. Studies have shown GAS pharyngitis outbreaks differentially infecting students at different schools within a town (31). Despite these limitations, we felt that pharyngitis retail health data provided the best available data to address the question of whether local incidence could help improve management for symptomatic patients.

## Conclusions

Since GAS pharyngitis occurs throughout the calendar year, live biosurveillance becomes particularly helpful. And since GAS pharyngitis is a common condition affecting millions per year in the US alone, small improvements in diagnostic accuracy have substantial impact. Incorporating real-time, biosurveillance-derived epidemiology into a prediction rule based on a patient's signs, symptoms, location and timing of presentation (32–34) suggests the value of this epidemiological analog to personalized genomic medicine (35), in which the context, instead of being derived from one's genes, derives from a quantitative representation of her epidemiologic milieu.

Developers of future recommendations for managing GAS pharyngitis should consider incorporating real time epidemiology with clinical factors. This novel approach in clinical guideline creation could help re-stratify risk for patients when the contemporaneous incidence of disease is high or very low, improving diagnostic accuracy. While our study focused on GAS pharyngitis, the results may have broad implications for other communicable diseases, where real-time biosurveillance data might be used to more accurately deduce the likelihood that a symptomatic individual has a specific disease.

The massive federal investment in health information technology (36) and the maturation of real-time biosurveillance systems present new opportunities to apply real-time epidemiology to individual patients (7). The $48 billion dollar investment is intended to support a "learning health system"(37) with bidirectional communication between clinical and public health sites, and delivery of clinical decision support to the point of care. Our findings

suggest that this learning health system should be capacitated to link live biosurveillance data with clinical decision-making.

## Acknowledgments

We thank CVS/Caremark and MinuteClinic for use of the data.

**Grant Support:** This work was supported by the Mentored Public Health Research Scientist Development Award K01 HK000055 from the Centers for Disease Control and Prevention (CDC), by Public Health Informatics Center of Excellence Award P01HK000088 from CDC, and by 1G08LM009778 and R01 LM007677 from the National Library of Medicine, National Institutes of Health. The study sponsors had no role in study design, collection, analysis or interpretation of data, writing of the manuscript or the decision to submit the manuscript for publication.

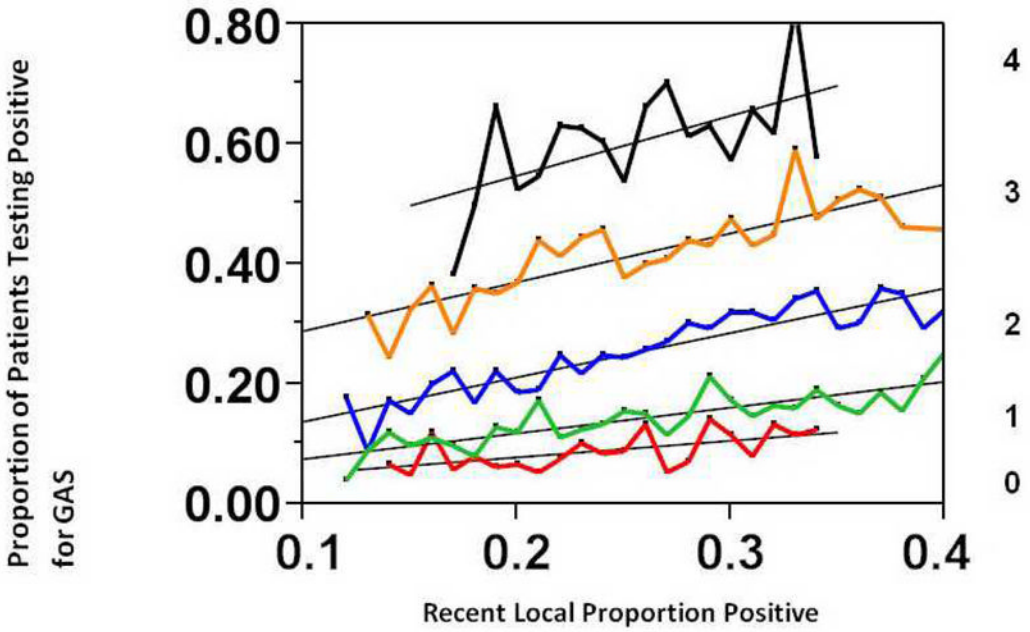
AMF had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## References

1. Fisman DN. Seasonality of infectious diseases. Annu Rev Public Health. 2007; 28:127–43. [PubMed: 17222079]
2. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. Am J Epidemiol. 2006; 164(10): 936–44. [PubMed: 16968863]
3. Syndromic surveillance. Reports from a national conference, 2004. MMWR Morb Mortal Wkly Rep. 2005; 54 (Suppl):1–180. [PubMed: 15647722]
4. Brownstein JS, Freifeld CC, Chan EH, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. N Engl J Med. 2010; 362(18):1731–5. [PubMed: 20445186]
5. Brownstein JS, Kleinman KP, Mandl KD. Identifying pediatric age groups for influenza vaccination using a real-time regional surveillance system. Am J Epidemiol. 2005; 162(7):686–93. [PubMed: 16107568]
6. Lipsitch M, Hayden FG, Cowling BJ, Leung GM. How to maintain surveillance for novel influenza A H1N1 when there are too many cases to count. Lancet. 2009; 374(9696):1209–11. [PubMed: 19679345]
7. Mandl KD, Overhage JM, Wagner MM, et al. Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience. J Am Med Inform Assoc. 2004; 11(2):141–150. [PubMed: 14633933]
8. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? Jama. 2005; 293(8):987–97. [PubMed: 15728170]
9. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. Lancet Infect Dis. 2005; 5(11):685–94. [PubMed: 16253886]
10. Ebell MH, Smith MA, Barry HC, Ives K, Carey M. The rational clinical examination. Does this patient have strep throat? Jama. 2000; 284(22):2912–8. [PubMed: 11147989]
11. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Med Decis Making. 1981; 1(3):239–46. [PubMed: 6763125]
12. Snow V, Mottur-Pilson C, Cooper RJ, Hoffman JR. Principles of appropriate antibiotic use for acute pharyngitis in adults. Ann Intern Med. 2001; 134(6):506–8. [PubMed: 11255529]
13. Bisno AL, Peter GS, Kaplan EL. Diagnosis of strep throat in adults: are clinical criteria really good enough? Clin Infect Dis. 2002; 35(2):126–9. [PubMed: 12087517]
14. McIsaac WJ, Kellner JD, Aufricht P, Vanjaka A, Low DE. Empirical validation of guidelines for the management of pharyngitis in children and adults. Jama. 2004; 291(13):1587–95. [PubMed: 15069046]
15. Asteberg I, Andersson Y, Dotevall L, et al. A food-borne streptococcal sore throat outbreak in a small community. Scand J Infect Dis. 2006; 38(11–12):988–94. [PubMed: 17148066]
16. Bisno AL. Acute pharyngitis. N Engl J Med. 2001; 344(3):205–11. [PubMed: 11172144]

17. Danchin MH, Rogers S, Kelpie L, et al. Burden of acute sore throat and group A streptococcal pharyngitis in school-aged children and their families in Australia. Pediatrics. 2007; 120(5):950–7. [PubMed: 17974731]
18. Martin JM, Green M, Barbadora KA, Wald ER. Erythromycin-resistant group A streptococci in schoolchildren in Pittsburgh. N Engl J Med. 2002; 346(16):1200–6. [PubMed: 11961148]
19. Kaplan EL, Wotton JT, Johnson DR. Dynamic epidemiology of group A streptococcal serotypes associated with pharyngitis. Lancet. 2001; 358(9290):1334–7. [PubMed: 11684215]
20. Mehrotra A, Liu H, Adams JL, et al. Comparing costs and quality of care at retail clinics with that of other medical settings for 3 common illnesses. Ann Intern Med. 2009; 151(5):321–8. [PubMed: 19721020]
21. Rudavsky R, Pollack CE, Mehrotra A. The geographic distribution, ownership, prices, and scope of practice at retail clinics. Ann Intern Med. 2009; 151(5):315–20. [PubMed: 19721019]
22. Pollack CE, Armstrong K. The geographic accessibility of retail clinics for underserved populations. Arch Intern Med. 2009; 169(10):945–9. discussion 950–3. [PubMed: 19468086]
23. Costello D. A checkup for retail medicine. Health Aff (Millwood). 2008; 27(5):1299–303. [PubMed: 18780914]
24. Institute for Clinical Systems Improvement Health Care Guideline. [Accessed April 29, 2011] Strep Pharyngitis Algorithm. http://www.icsi.org/respiratory_illness_in_children_and_adults_guideline_/respiratory_illness_in_children_and_adults__guideline__13116.html
25. Woodburn JD, Smith KL, Nelson GD. Quality of care in the retail health care setting using national clinical guidelines for acute pharyngitis. Am J Med Qual. 2007; 22(6):457–62. [PubMed: 18006426]
26. Corneli HM. Rapid strep tests in the emergency department: an evidence-based approach. Pediatr Emerg Care. 2001; 17(4):272–8. quiz 279. [PubMed: 11493831]
27. Middleton KR, Hing E. National Hospital Ambulatory Medical Care Survey: 2004 outpatient department summary. Adv Data. 2006; (373):1–27.
28. Wigton RS, Connor JL, Centor RM. Transportability of a decision rule for the diagnosis of streptococcal pharyngitis. Arch Intern Med. 1986; 146(1):81–3. [PubMed: 3510600]
29. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003; 78(8):775–80. [PubMed: 12915363]
30. Kahneman, D.; Slovic, P.; Tversky, A. Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press; 1982.
31. Cornfeld D, Hubbard JP, Harris TN, Weaver R. Epidemiologic studies of streptococcal infection in school children. Am J Public Health Nations Health. 1961; 51:242–9. [PubMed: 13695599]
32. Fine A, Brownstein J, Nigrovic L, et al. Integrating Spatial Epidemiology Into a Decision Model for Evaluation of Facial Palsy in Children. Arch Pediatr Adolesc Med. 2011; 165(1):61–67. [PubMed: 21199982]
33. Fine A, Reis B, Nigrovic L, et al. Use of Population Health Data to Refine Diagnostic Decision Making for Pertussis. J Am Med Inform Assoc. 2010; 17(1):85–90. [PubMed: 20064807]
34. Fine AM, Nigrovic LE, Reis BY, Cook EF, Mandl KD. Linking Surveillance to Action: Incorporation of Real-time Regional Data into a Medical Decision Rule. J Am Med Inform Assoc. 2007; 14(2):206–211. [PubMed: 17213492]
35. Snyderman R, Dinon MA. Improving Health by Taking It Personally. JAMA. 2010; 303(4):363–364. [PubMed: 20103762]
36. Blumenthal D. Launching HITECH. N Engl J Med. 2010; 362(5):382–5. [PubMed: 20042745]
37. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. Sci Transl Med. 2010; 2(57):57cm29.

## Appendix




**Appendix Figure 3. Proportion of patients in the Validation set (n=27,081) testing positive for Group A Streptococcal pharyngitis by recent local proportion positive (RLPP) and grouped and labeled by Centor score**

Each line represents patients with the same Centor score across varying RLPP. The numbers at the far right side of the graph identify the Centor score for each line. The proportion of patients who tested positive increases both as the clinical score increases, and as the RLPP increases. Pearson's coefficient was used to measure strength of correlation. The $r^2$, representing the proportion of the variation in GAS pharyngitis positivity that can be attributed to the RLPP, is 0.33, (p=0.012) for Centor 4, 0.70 (p<0.001) for Centor 3, 0.82 (p<0.001) for Centor 2, 0.68 (p<0.001) for Centor 1, and 0.35 (p=0.005) for Centor 0. The slopes of the lines for Centor 4, 3, 2, 1, and 0 are 0.99, 0.80, 0.75, 0.43 and 0.28. Each data point represents a median of 140 patients (range 45–555 patients, interquartile range 82–290).

### Appendix – Table 4

Reclassification Accuracy of *Adjusted* Centor Score Resulting From Incrementing Score by One Point at Different Thresholds of GAS Pharyngitis Activity (Validation set)

| | | Number reclassified GAS pharyngitis positive | | | Number reclassified per 1000 | | National estimates of numbers affected | |
|---|---|---|---|---|---|---|---|---|
| **Centor** | **RLPP threshold** | **Incorrect** | **Correct** | **Net (%)** | **Incorrect** | **Correct** | **Incorrect** | **Correct** |
| **1** | >0.20 ($n_1$=8681) | 373 | 974 | 601 | 43 | 112 | 63,118 | 164,401 |
| | >0.25 ($n_1$=3827) | 160 | 504 | 344 | 42 | 132 | 39,543 | 124,277 |
| | >0.30 ($n_1$=1734) | 72 | 236 | 164 | 41 | 136 | 18,103 | 60,048 |
| | >0.35 ($n_1$=633) | 26 | 91 | 65 (10) | 41 | 144 | 7171 | 25,186 |
| **2** | >0.20 ($n_2$=9378) | 9738 | 747 | −8990 (−64) | 710 | 50 | 1,125,725 | 79,276 |

| | | Number reclassified GAS pharyngitis positive | | | Number reclassified per 1000 | | National estimates of numbers affected | |
|---|---|---|---|---|---|---|---|---|
| **Centor** | **RLPP threshold** | **Incorrect** | **Correct** | **Net (%)** | **Incorrect** | **Correct** | **Incorrect** | **Correct** |
| | >0.25 ($n_2$=4145) | 6555 | 537 | −6018 (−63) | 657 | 62 | 668,146 | 63,052 |
| | >0.30 ($n_2$=1969) | 3117 | 293 | −2824 (−60) | 635 | 66 | 302,485 | 19,988 |
| | >0.35 ($n_2$=772) | 1150 | 122 | −1028 (−57) | 630 | 67 | 119,019 | 12,658 |

Adjustment of Centor Score to reclassify risk according to prior probability of disease as inferred by the recent local proportion positive (RLPP). For patients with a Centor score of 1 or 2, the score was incremented by one when RLPP exceeded thresholds. Changing from 1 to 2 would result in testing $n_1$ patients for GAS pharyngitis, and changing from 2 to 3 would result in treating $n_2$ patients empirically. Net change (correct minus incorrect) and net percent change are shown. The number reclassified correctly and incorrectly standardized per 1000 patients, and the national estimates of the numbers affected are shown.

## Appendix Table 5

Reclassification Accuracy of *Adjusted* Centor Score Resulting From Decreasing Score by One Point at Different Thresholds of GAS Pharyngitis Activity (Validation set)

| | | Number reclassified GAS pharyngitis negative | | | Number reclassified per 1000 | | National estimates of numbers affected | |
|---|---|---|---|---|---|---|---|---|
| **Centor** | **RLPP threshold** | **Incorrect** | **Correct** | **Net (%)** | **Incorrect** | **Correct** | **Incorrect** | **Correct** |
| **2** | RLPP< 0.25 ($n_2$=4395) | 722 | 175 | −548 (−12) | 164 | 40 | 169,260 | 41,283 |
| | RLPP <0.20 ($n_2$=2179) | 330 | 88 | −242 (−11) | 152 | 41 | 75,765 | 20,436 |
| | RLPP < 0.15 ($n_2$=388) | 49 | 16 | −33 (−8) | 126 | 42 | 11,391 | 3797 |
| **3** | RLPP < 0.25 ($n_3$=2169) | 165 | 1275 | 1110 | 76 | 588 | 42,039 | 325,246 |
| | RLPP < 0.20 ($n_3$=1044) | 69 | 663 | 594 (57) | 66 | 635 | 17,632 | 169,637 |
| | RLPP < 0.15 ($n_3$=201) | 12 | 134 | 122 (61) | 60 | 666 | 2907 | 32,268 |

Adjustment of Centor Score to reclassify risk according to prior probability of disease as inferred by the recent local proportion positive (RLPP). For patients with a Centor score of 2 or 3, the score was decreased by one when RLPP was below thresholds. Changing from 2 to 1 would result in not testing or treating $n_2$ patients for GAS pharyngitis, and changing from 3 to 2 would result in testing $n_3$ patients and treating them if they tested positive. Net change (correct minus incorrect) and net percent change are shown. The number reclassified correctly and incorrectly standardized per 1000 patients, and the national estimates of the numbers affected are shown.

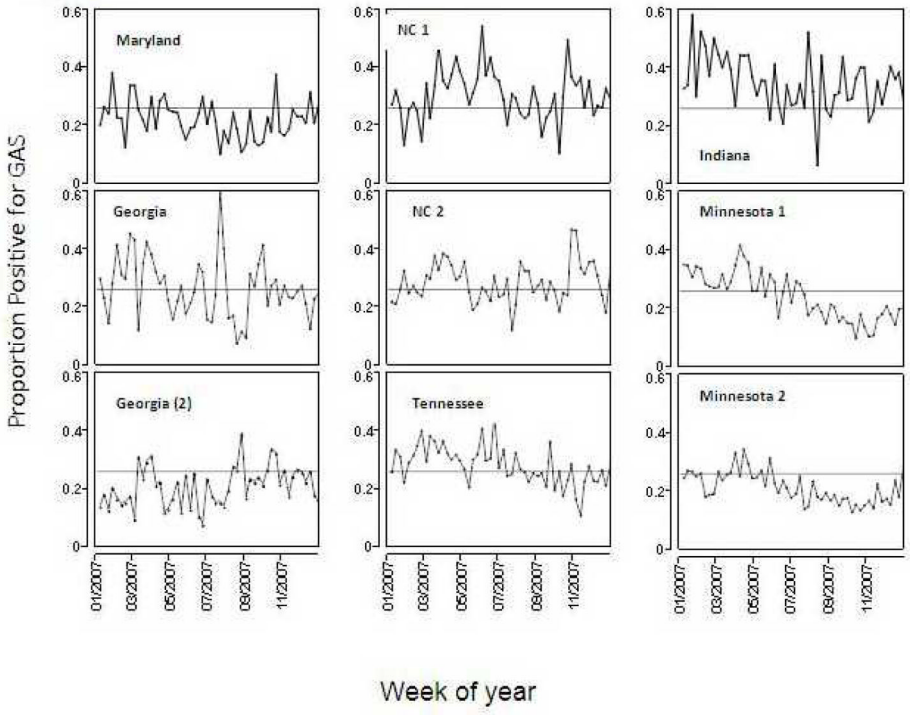



**Figure 1. Proportion positive by study week for nine different locations**

The *x* axis is the study week from January 1, 2007– December 31, 2007, and the *y* axis is weekly proportion positive for group A streptococcal pharyngitis. Each graph shows the proportion of patients who tested positive each week in one of nine markets. The axes have been standardized to allow comparison across markets. The horizontal line is the average across all markets (0.25), and is provided for reference and to facilitate comparison.

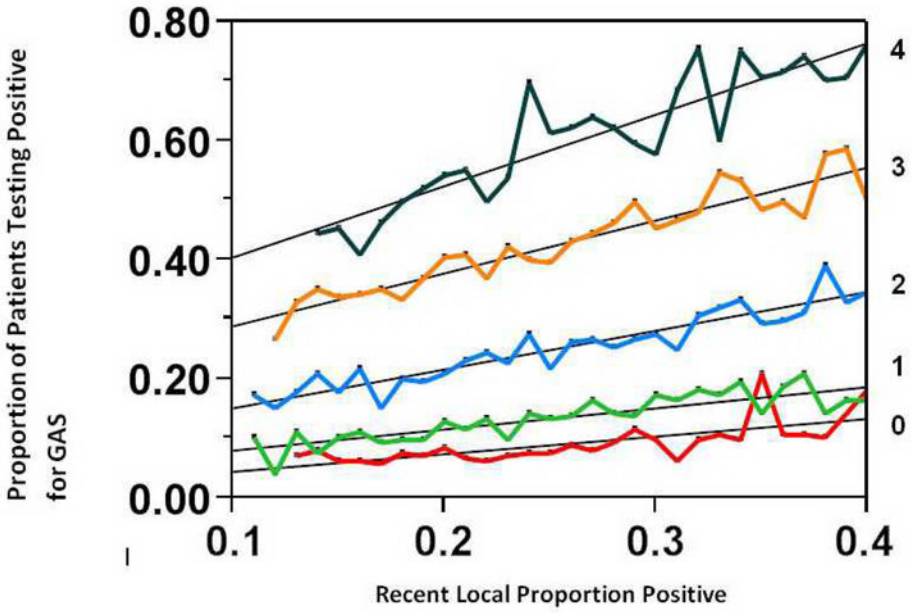



**Figure 2. Proportion of patients testing positive for Group A Streptococcal pharyngitis by recent local proportion positive (RLPP) and grouped and labeled by Centor score**

Each line represents patients with the same Centor score across varying RLPP. The numbers at the far right side of the graph identify the Centor score for each line. The proportion of patients who tested positive increases both as the clinical score increases, and as the RLPP increases. Pearson's coefficient was used to measure strength of correlation. The $r^2$, representing the proportion of the variation in GAS pharyngitis positivity that can be attributed to the RLPP, is 0.81, ($p<0.001$) for Centor 4, 0.86 ($p<0.001$) for Centor 3, 0.86 ($p<0.001$) for Centor 2, 0.70 ($p<0.001$) for Centor 1, and 0.47 ($p<0.001$) for Centor 0. The slopes of the lines for Centor 4, 3, 2, 1, and 0 are 1.21, 0.88, 0.64, 0.37 and 0.29. Each data point represents a median of 223 patients (range 41–1152 patients, interquartile range 115–518).

## Table 1

Clinical characteristics of patients presenting with pharyngitis to the retail health clinics (derivation set: n=54,981 patient-visits)

| Characteristic | Overall (n=54,981) | Group A Streptococcal positive (n=13,823) | Group A Streptococcal negative (n=41,158) |
|---|---|---|---|
| Female (%) | 68 | 66 | 68 |
| Mean age in years (Median, IQ range) | 33.5 (33, 24–41) | 33.2 (33, 25–40) | 33.5 (33, 24–41) |
| Age group (%)<br>15–18 years | 13 | 10 | 14 |
| 19–39 years | 58 | 64 | 56 |
| 40–59 years | 26 | 24 | 26 |
| >60 years | 3 | 2 | 3 |
| Fever (%) | 31 | 45 | 26 |
| Swollen anterior cervical lymph nodes (%) | 61 | 77 | 55 |
| Absence of cough (%) | 68 | 76 | 66 |
| Tonsillar exudates (%) | 22 | 40 | 16 |
| Distribution by score<br>Centor 0 (%) | 8 | 3 | 10 |
| Centor 1 (%) | 33 | 18 | 38 |
| Centor 2 (%) | 34 | 34 | 35 |
| Centor 3 (%) | 18 | 31 | 14 |
| Centor 4 (%) | 6 | 15 | 3 |

IQ : interquartile

To calculate the Centor score for a patient who presents with acute pharyngitis, one point is given for each of the following items: presence of tonsillar exudates, history of fever, presence of swollen anterior cervical lymph nodes, and absence of cough (11).

**Table 2**

Reclassification Accuracy of *Adjusted* Centor Score Resulting From Incrementing Score by One Point at Different Thresholds of GAS Pharyngitis Activity

| | | Number reclassified GAS pharyngitis positive | | | Number reclassified per 1000 | | National estimates of numbers affected | |
|---|---|---|---|---|---|---|---|---|
| **Centor** | **RLPP threshold** | **Incorrect** | **Correct** | **Net (%)** | **Incorrect** | **Correct** | **Incorrect** | **Correct** |
| **1** | >0.20 ($n_1$=13,056) | 556 | 1542 | 985 (8) | 43 | 118 | 63,647 | 174,658 |
| | >0.25 ($n_1$=8796) | 370 | 1119 | 749 | 42 | 127 | 40,058 | 121,129 |
| | >0.30 ($n_1$=4355) | 180 | 607 | 427 (10) | 41 | 139 | 18,446 | 62,537 |
| | >0.35 ($n_1$=1579) | 65 | 216 | 151 (10) | 41 | 137 | 7394 | 24,707 |
| **2** | >0.20 ($n_2$=13,987) | 9738 | 747 | −8990 (−64) | 696 | 53 | 1,087,646 | 82,824 |
| | >0.25 ($n_2$=9583) | 6555 | 537 | −6018 (−63) | 684 | 56 | 688,768 | 56,390 |
| | >0.30 ($n_2$=4746) | 3117 | 293 | −2824 (−60) | 657 | 62 | 312,077 | 29,450 |
| | >0.35 ($n_2$=1818) | 1150 | 122 | −1028 (−57) | 632 | 67 | 120,336 | 12,757 |

Adjustment of Centor Score to reclassify risk according to prior probability of disease as inferred by the recent local proportion positive (RLPP). For patients with a Centor score of 1 or 2, the score was incremented by one when RLPP exceeded thresholds. Changing from 1 to 2 would result in testing $n_1$ patients for GAS pharyngitis, and changing from 2 to 3 would result in treating $n_2$ patients empirically. Net change (correct minus incorrect) and net percent change are shown. The number reclassified correctly and incorrectly standardized per 1000 patients, and the national estimates of the numbers affected are shown.

*Ann Intern Med*. Author manuscript; available in PMC 2013 May 12.

**Table 3**

Reclassification Accuracy of *Adjusted* Centor Score Resulting From Decreasing Score by One Point at Different Thresholds of GAS Pharyngitis Activity

| | | Number reclassified GAS pharyngitis negative | | | Number reclassified per 1000 | | National estimates of numbers affected | |
|---|---|---|---|---|---|---|---|---|
| **Centor** | **RLPP threshold** | **Incorrect** | **Correct** | **Net (%)** | **Incorrect** | **Correct** | **Incorrect** | **Correct** |
| **2** | RLPP< 0.25 ($n_2$=8746) | 1498 | 344 | −1155 (−13) | 171 | 39 | 175,950 | 40,129 |
| | RLPP <0.20 ($n_2$=4342) | 655 | 176 | −479 (−11) | 151 | 41 | 76,433 | 20,753 |
| | RLPP < 0.15 ($n_2$=847) | 126 | 35 | −91 (−11) | 148 | 41 | 13,741 | 3807 |
| **3** | RLPP < 0.25 ($n_3$=4334) | 325 | 2574 | 2249 (52) | 75 | 594 | 40,971 | 324,492 |
| | RLPP < 0.20 ($n_3$=2080) | 145 | 1289 | 1145 (55) | 70 | 620 | 18,812 | 166,616 |
| | RLPP < 0.15 ($n_3$=371) | 25 | 233 | 208 (56) | 68 | 627 | 3352 | 30,907 |

Adjustment of Centor Score to reclassify risk according to prior probability of disease as inferred by the recent local proportion positive (RLPP). For patients with a Centor score of 2 or 3, the score was decreased by one when RLPP was below thresholds. Changing from 2 to 1 would result in not testing or treating $n_2$ patients for GAS pharyngitis, and changing from 3 to 2 would result in testing $n_3$ patients and treating them if they tested positive. Net change (correct minus incorrect) and net percent change are shown. The number reclassified correctly and incorrectly standardized per 1000 patients, and the national estimates of the numbers affected are shown.

*Ann Intern Med*. Author manuscript; available in PMC 2013 May 12.