Performance of Rapid Streptococcal Antigen TestingVaries by Personnel

James W. Fox,1* Daniel M. Cohen,2 Mario J. Marcon,3 William H. Cotton,4 and Bema K. Bonsu2

Department of Pediatrics, Division of Emergency Medicine, Children’s Hospital of Akron, Akron, Ohio,1 and Department of Pediatrics, Divisions of Emergency Medicine,2 Laboratory Medicine,3 and Ambulatory Pediatrics,4 Children’s Hospital, Columbus, Ohio

Received 6 July 2006/Returned for modification 30 July 2006/Accepted 30 August 2006

Rapid carbohydrate antigen tests are frequently used to diagnose group A streptococcal (GAS) pharyngitis. Despite evidence of modest sensitivity in medical settings, rapid antigen tests are available to the public for self-testing. We sought to determine if the personnel performing a rapid streptococcal antigen test influence the test’s performance characteristics. Throat swabs of pediatric patients performed for GAS pharyngitis in a tertiary-care children’s hospital network were included during two study periods in 2004 and 2005. The performance characteristics of a rapid carbohydrate antigen test were evaluated in three clinical settings against a nucleic acid probe test method according to the personnel performing the test (laboratory technologist versus nonlaboratory personnel). Between the study periods, nonlaboratory personnel from one site underwent retraining. Subsequently, the performance characteristics of the rapid antigen test were reassessed. The sensitivity of the rapid antigen test varied widely among the different testing sites (56 to 90%). Notably, test sensitivity was consistently greater when the test was performed by laboratory technologists than when it was performed by nonlaboratory personnel (P < 0.0001). Although the rapid antigen test sensitivity significantly improved after nonlaboratory personnel at one testing site were retrained (sensitivity before versus after retraining; P < 0.0001), the sensitivity remained greater in the laboratory technologist cohort (P < 0.0001). These data confirm the important relationship of the operator performing a rapid streptococcal antigen test with the test’s accuracy, even in a clinical setting, where operator training is mandated. Therefore, its use outside the medical setting by lay persons cannot be recommended without culture backup.

Group A streptococcus (GAS) is a common cause of pharyngitis, resulting in more than 10 million physician visits each year (6) and accounting for 15 to 30% of sore throats in children (3, 11, 18, 29). Accurate diagnosis is necessary to permit targeted administration of antimicrobial therapy to prevent suppurrative (peritonsillar and retropharyngeal abscesses) and nonsuppurrative (rheumatic heart disease and acute glomerulonephritis) complications, to hasten symptom resolution, and to reduce the transmission of GAS in the community while limiting the use of antibiotics in virus-mediated infections (10, 22, 26, 32, 36).

Unfortunately, clinical findings and simple scoring systems based on these findings are unreliable for identifying patients with GAS pharyngitis (2, 12, 17, 24, 31). For this reason, laboratory tests are commonly used to confirm the diagnosis in children who present with throat pain. Rapid streptococcal carbohydrate antigen tests are widely used in the clinical setting because the results are available during a patient’s visit. The ability to provide treatment at the point of care is a distinct advantage of this test compared to standard culture. Importantly, a primary drawback of carbohydrate detection systems is modest sensitivity (55 to 90% in postmarketing trials) (1, 7, 14, 16, 19, 23, 30, 34). Therefore, the American Academy of Pediatrics (8), the American Heart Association (9), and the Infectious Disease Society of America (4) all recommend backup throat cultures for negative rapid antigen detection tests in pediatric patients to achieve greater overall diagnostic sensitivity.

Although the sale of rapid antigen test kits directly to individuals for use in the home setting is not approved by the United States Food and Drug Administration (FDA), select kits may be purchased in a pharmacy or over the Internet (27, 28). Thus, these testing devices are being utilized without physician supervision and without the safeguard of a follow-up culture. Furthermore, while a positive antigen test result for GAS is likely to prompt consultation with a physician, a negative test result may falsely imply illness due to another etiologic agent, resulting in a delayed diagnosis or no diagnosis at all. Regrettably, as self-testing kits become increasingly available to the public, undiagnosed and untreated GAS infections are likely to increase, serving as a reservoir for the spread of the infection to the community and putting patients with these infections at risk for complications of streptococcal tonsillopharyngitis.

A number of factors have the potential to further lower the sensitivity of antigen tests performed by patients in the home setting. These factors include improper collection of samples from sites other than the pharynx or tonsils, insufficient quantity of sample, and inability to perform and interpret the test correctly. All but the last of these factors are likely to be overcome for a cooperative child by adequately instructing parents in proper techniques of specimen collection. Interpretation of the test, however, is likely to reflect the operator’s experience in performing and reading the test. If so, the accu-
racy of GAS antigen tests may be relatively low when the tests are performed by nonlaboratory personnel compared to laboratory personnel.

In this study, we evaluated the impacts of education, training, and experience on the performance of a rapid streptococcal antigen detection test. Specifically, we gauged the utility of testing in the home setting by comparing the accuracy of a simple antigen test when performed by laboratory versus nonlaboratory personnel at a tertiary-care children’s hospital and associated satellite urgent-care facilities. (The rapid streptococcal antigen test studied here is no longer available through Abbott Laboratories. The same test is now sold by Genzyme Diagnostics [San Diego, CA] as the Genzyme OSOM Strep A test.)

MATERIALS AND METHODS

Study population. Pediatric patients being evaluated for streptococcal pharyngitis in the Columbus Children’s Hospital medical system were included in this study. This medical care system consists of a tertiary-care emergency department, four urgent care clinics, and 10 primary care centers. Annually, nearly 500,000 patient visits are recorded at these sites. Patients were included if a throat swab for the detection of group A streptococcus was performed. The study was divided into two periods. Period 1 included patients evaluated at any of the sites during October and November 2004, while period 2 included patients evaluated during January 2005 to 29 January 2005; 26 days), nonlaboratory personnel performing rapid antigen testing in the primary care centers underwent competency assessment and retraining by microbiology laboratory personnel.

Sample collection and testing. A Dacron double-swab collection-transport system (COPAN Venturi Transport; COPAN Diagnostics Inc., Corona, CA) was used for all sampling. Samples were collected from the patient’s posterior pharynx and tonsillar surfaces by physicians and nurses as recommended by the Infectious Disease Society of America (4). Immediately after sample collection, one swab from each dual-swab collection-transport device was used for rapid antigen detection testing (Signify Strep A; Abbott Laboratories, Abbott Park, IL) at all locations. The remaining swab was stored in a secure refrigerator and transported to the Columbus Children’s Hospital Research Institute of Columbus Children’s Hospital. Because routine patient care was not altered and patient data were not linked to test results during data analysis, informed consent was waived.

RESULTS

Tables 1 and 2 show the test characteristics of the rapid antigen test for study period 1 (Table 1) and study period 2 (Table 2). While the test specificity was excellent (≥97%) at all sites for both time periods, the sensitivity varied from 56 to 90%. The sensitivity of the rapid antigen test prior to retraining (study period 1) did not differ between the nonlaboratory personnel groups (P = 0.29). However, the sensitivity of the test when performed by laboratory personnel was significantly greater than the sensitivity observed for testing by nonlaboratory personnel (P < 0.0001 for pairwise comparisons) (Table 1).

In the second period, the sensitivity of the rapid antigen test improved significantly after competency retraining in the primary-care centers (“Nonlab, retrained”; P < 0.0001). However, a significantly larger improvement was observed in nonlaboratory personnel who did not undergo retraining (an absolute increase in sensitivity of 26% versus 15% for “Nonlab, control” and “Nonlab, retrained,” respectively; P = 0.016).

**TABLE 1. Performance characteristics of rapid antigen testing by personnel prior to the retraining of a nonlaboratory cohort, period 1**

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of tests run (n)</th>
<th>Prevalence (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonlab, retraineda</td>
<td>261</td>
<td>16</td>
<td>60b,c</td>
<td>99</td>
</tr>
<tr>
<td>Nonlab, controlb</td>
<td>497</td>
<td>16</td>
<td>56b</td>
<td>98</td>
</tr>
<tr>
<td>Labc</td>
<td>434</td>
<td>22</td>
<td>88</td>
<td>99</td>
</tr>
</tbody>
</table>

a Retraining before period 2.
b P < 0.0001 (all P values were calculated by the binomial test for proportions); pairwise comparison to “Lab” group.
c P > 0.05; pairwise comparison to “Nonlab, control” group.
d No retraining before period 2.

**TABLE 2. Performance characteristics of rapid antigen testing by personnel after the retraining of a nonlaboratory cohort, period 2**

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of tests run (n)</th>
<th>Prevalence (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonlab, retraineda</td>
<td>377</td>
<td>25</td>
<td>75b,c</td>
<td>97</td>
</tr>
<tr>
<td>Nonlab, controlb</td>
<td>843</td>
<td>20</td>
<td>82b</td>
<td>98</td>
</tr>
<tr>
<td>Labc</td>
<td>475</td>
<td>31</td>
<td>90</td>
<td>97</td>
</tr>
</tbody>
</table>

a Retraining before period 2.
b P < 0.05 (all P values were calculated by the binomial test for proportions); pairwise comparison to “Nonlab, control” group.
c P < 0.0001; pairwise comparison to “Lab” group.
d No retraining before period 2.
Despite the improvement in the sensitivity of the rapid antigen test performed by nonlaboratory personnel in the second study period, the sensitivity of the GAS tests remained superior when performed by laboratory personnel ($P < 0.0001$ for pairwise comparisons) (Table 2).

The number of tests run in the second period by nonlaboratory personnel who did not undergo retraining (843) was over two times greater than the number run in the same period by nonlaboratory personnel who underwent retraining (377). Additionally, the prevalence was significantly higher in study period 2 than in study period 1 ($P < 0.001$).

**DISCUSSION**

Many variables affect the performance of laboratory tests. For a rapid streptococcal carbohydrate antigen detection system, in particular, predictors include the quality of the specimen obtained from the posterior pharynx and tonsillar tissue, as it relates to the inoculum size, and the operator’s ability to appropriately test the throat sample and interpret the result of the test (12, 13, 15, 21, 23). Data from our study confirm that the operator performing and interpreting the test is an important determinant of the test’s accuracy, even in a clinical setting, where training and quality assurance are mandated. Specifically, our data show that even when dealing with a relatively simple test, in a hospital system that cares for nearly 500,000 pediatric patients annually, performance varies widely based on the experience of the operator. Additionally, our test operators were medical personnel who were likely to be more cognizant of proper kit storage, expiration dates, and use of controls than the lay public. In extrapolating these results to the performance of the test when used for self-diagnosis in the home setting, our data indicate that the accuracy of the test is likely to be lower.

A number of factors may play a role in explaining the observed differences in test accuracy between laboratory and nonlaboratory personnel. We believe, however, that the most important determinant of accuracy over time is accumulated operator experience. This may merely reflect the number of tests run by an operator (as suggested by our data for nonlaboratory personnel) or echo qualitative factors, such as whether feedback is provided routinely concerning the accuracy of putative diagnoses when the test is performed. Specifically, if there are regular opportunities (particularly in borderline cases) for test operators to observe features of screening tests that reflect a higher likelihood of a positive bacterial-culture result, these observations are likely to enhance learning. Such ideal conditions for learning are likely to be present in settings with full laboratory support, where both antigen testing and culture are often performed by the same operator(s), but are less probable in settings where only the initial antigen test is performed without regular culture feedback, as is likely to occur in settings that lack full on-site laboratory support. Furthermore, conditions such as these—high frequency of testing and performance enhancement from subsequent culture results—would not be present in the home setting. Therefore, significant improvement in testing performance would not be expected for the lay user.

Another factor that may play a part in differences in test performance is the severity of the streptococcal pharyngitis. Indeed, studies have found that the sensitivity of rapid streptococcal antigen tests improves as the clinical likelihood of streptococcal pharyngitis increases (spectrum bias) (11, 12, 18). In our study, to the extent that the severity of illness was related to the site of the clinical encounter and confounded by the type/experience of personnel running tests at these sites, the difference in the accuracy of tests may have been influenced by spectrum bias. In fact, the sensitivity improved for each testing site during the second study period, which showed a higher prevalence of disease (GAS pharyngitis). Nonetheless, in extrapolating this factor to self-testing outside of the medical setting, it is reasonable to expect that rapid antigen tests, if performed more often in children with less severe symptoms—a likely scenario—will perform as poorly due to the same sort of bias.

Interestingly, we observed that rapid antigen test sensitivity improved significantly for both nonlaboratory personnel groups (Table 2), even though only one nonlaboratory group underwent retraining between study periods 1 and 2. This finding may be related to the increased rate of GAS testing during the second study period, which provided increased experience from performing the antigen test more frequently. Tests were conducted more frequently among nonlaboratory operators who were not undergoing retraining and appeared to have a greater impact on test accuracy than retraining. Thus, it is likely that retraining did play a role in the improvement in sensitivity seen in the experimental group but that the effects were masked by the improvement seen in all groups. However, without quality improvement measures, it is highly improbable that the sensitivity of the test for self-diagnosis by patients would approach even the low rates found in our study (sensitivity as low as 56%).

In addition, the improved sensitivity for the nonlaboratory groups may be explained by the “Hawthorne effect” (25). This theory suggests that performance may improve, often transiently, simply because the participant knows he or she is being observed. For our study, between the study periods, operators may have become aware that the performance of the rapid strep test was under evaluation. Therefore, the operators may have used more care in performing the tests during study period 2, thus affecting the results of the study.

Until the major determinants of test performance accuracy are elucidated and controlled, it appears imprudent to promote home GAS antigen kits for self-diagnosis. We concur with recommendations by the American Academy of Pediatrics, the American Heart Association, and the Infectious Disease Society of America, which recommend backup testing for negative rapid streptococcal antigen tests in children—testing that is presently available only in the clinical setting. The FDA has enforced these recommendations: because “no rapid [streptococcal] test has been cleared, approved, or waived through the regulatory process as a stand alone [sic] test in the face of locally suppurative disease, lack of a backup method for a negative rapid GAS test result constitutes off label [sic] use” (35).

Notably, even with the best sensitivity of 90% measured in this study, 1 in 10 children with a negative rapid antigen result would be at risk for suppurative and nonsuppurative complications of streptococcal pharyngitis without culture confirmation and subsequent treatment. It is relevant to note that, when
performing the rapid streptococcal antigen test, 76% of parents state that they would not seek medical evaluation if the test was negative (5). With the potential for widening use of the home kits in the absence of backup cultures, a large number of such patients are unlikely to be recognized as having GAS pharyngitis and could pose a public health risk.

Our study has several limitations. A number of factors that may have affected the outcome of this investigation were not studied. They included differences in collection techniques for obtaining throat swab specimens, differences in the interval to performing the rapid antigen test, and differences in patient characteristics at the various sites during the study period. All of these factors may have acted as confounders when test performances at the different sites were compared. Future studies should evaluate the variability in test performance between paired laboratory and nonlaboratory personnel using the same specimens and after controlling for these factors. In addition, our study extrapolates findings for the test from the clinical setting to the home setting for patients who intend to or already utilize the rapid antigen test for self-diagnosis. Such guarded extrapolation of study findings is reasonable. However, future studies are needed to evaluate directly the performance characteristics of these rapid antigen tests against culture results in this setting. Finally, we investigated the performance of a single rapid antigen test. There are a number of rapid streptococcal antigen tests on the market, and it is possible that another test may perform differently than reported here. Future studies should be performed to investigate this potential discrepancy.

In conclusion, our study results suggest that the accuracy of rapid antigen tests for detecting GAS in throat specimens varies widely depending on the personnel performing the test, even in a clinical setting that mandates training and retraining; this ostensibly reflects the accumulated experience of test operators. It is pertinent that the accuracy of antigen tests when performed by laboratory-employed personnel consistently surpassed that of tests performed by nonlaboratory personnel, and among the latter, testing frequency was more important than retraining. Because rapid tests fail to exclude GAS pharyngitis in a modest fraction of patients, and because both positive and negative results mandate evaluation by a medical professional, we do not recommend the use of these tests for self-evaluation of throat pain in the home setting. Backup culture of negative rapid antigen tests should be performed in all pediatric settings to reliably diagnose GAS pharyngitis.

ACKNOWLEDGMENTS

We thank Soledad Fernandez and Namhee Kim for their statistical expertise in the preparation of the manuscript. We also thank Marilyn Hribar for her technical assistance.

There were no potential, perceived, or real conflicts of interest, financial or otherwise, in producing this study. There was no financial support for this investigation.

REFERENCES