Spectrum and Inoculum Size Effect of a Rapid Antigen Detection Test for Group A Streptococcus in Children with Pharyngitis

Jérémie F. Cohen1,2*, Martin Chalumeau1,2, Corinne Levy3, Philippe Bidet4, Franck Thollot5, Alain Wollner3, Edouard Bingen4, Robert Cohen3,6

1 Unité de Recherche Epidémiologique en Santé Périmatale et Santé des Femmes et des Enfants (U953), Inserm, Paris, France, 2 Service de Pédiatrie Générale, Hôpital Necker-Enfants-Malades, AP-HP et Université Paris-Descartes, Paris, France, 3 Association Clinique et Thérapeutique du Val-de-Marne (ACTIV), Saint-Maur-des-Fossés, France, 4 Service de Microbiologie, Hôpital Robert Debré, AP-HP et Université Paris-Diderot, Paris, France, 5 Association Française de Pédiatrie Ambulatoire (AFPA), Essey-lès-Nancy, France, 6 Service de Microbiologie, Centre Hospitalier Intercommunal de Créteil, Créteil, France

Abstract

Background: The stability of the accuracy of a diagnostic test is critical to whether clinicians can rely on its result. We aimed to assess whether the performance of a rapid antigen detection test (RADT) for group A streptococcus (GAS) is affected by the clinical spectrum and/or bacterial inoculum size.

Methods: Throat swabs were collected from 785 children with pharyngitis in an office-based, prospective, multicenter study (2009–2010). We analysed the effect of clinical spectrum (i.e., the McIsaac score and its components) and inoculum size (light or heavy GAS growth) on the accuracy (sensitivity, specificity, likelihood ratios and predictive values) of a RADT, with laboratory throat culture as the reference test. We also evaluated the accuracy of a McIsaac-score–based decision rule.

Results: GAS prevalence was 36% (95CI: 33%–40%). The inoculum was heavy for 85% of cases (81%–89%). We found a significant spectrum effect on sensitivity, specificity, likelihood ratios and positive predictive value (p<0.05) but not negative predictive value, which was stable at about 92%. RADT sensitivity was greater for children with heavy than light inoculum (95% vs. 40%, p<0.001). After stratification by inoculum size, the spectrum effect on RADT sensitivity was significant only in patients with light inoculum, on univariate and multivariate analysis. The McIsaac-score–based decision rule had 99% (97%–100%) sensitivity and 52% (48%–57%) specificity.

Conclusions: Variations in RADT sensitivity only occur in patients with light inocula. Because the spectrum effect does not affect the negative predictive value of the test, clinicians who want to rule out GAS can rely on negative RADT results regardless of clinical features if they accept that about 10% of children with negative RADT results will have a positive throat culture. However, such a policy is more acceptable in populations with very low incidence of complications of GAS infection.

Introduction

Group A streptococcus (GAS) is found in 20% to 40% of cases of childhood pharyngitis; the remaining cases are considered viral [1]. Clinical examination cannot distinguish accurately between viral and GAS pharyngitis [2], and a diagnostic test based on a throat swab is recommended in most countries [3]. Throat culture on a blood agar plate in a microbiology laboratory for 48 hours is the reference test for diagnosis of GAS pharyngitis [4–7]. Rapid antigen detection tests (RADTs) have been proposed as an alternative to throat culture. Compared to laboratory culture, RADTs have high specificity (≈95%) and results are immediate.

Their main drawbacks are low sensitivity (∼85%, range 65.6% to 98.9%) [8,9] and variations in sensitivity by clinical spectrum of the disease (spectrum effect, or spectrum bias) [10–12]. The major issue for the spectrum effect is the generalisability of test performance. First, the overall population estimate might not be generalisable to patient subpopulations; second, the diagnostic accuracy that was observed in one study might not be applicable to other patients. Three studies have shown a significant spectrum effect on RADT sensitivity, but these studies had methodological limitations suggesting selection, indication, partial verification, and measurement biases [13–15]. No studies investigated the effect of


Editor: Bernard Beall, Centers for Disease Control & Prevention, United States of America

Received March 7, 2012; Accepted May 18, 2012; Published June 29, 2012

Copyright: © 2012 Cohen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was funded by Dectrapharm, manufacturer of the RADT (Rapid Antigen Detection Test). JFC is a masters student supported by educational grants from Agence Régionale de Santé d’Île-de-France and Laboratoires Guigoz - Société Française de Pédiatrie - Groupe de Pédiatrie Générale - Groupe de Recherches Épidémiologiques en Pédiatrie. This work was supported by a research grant from the Assistance Publique-Hôpitaux de Paris (PHRC Regional 2012, AOR 12089). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: This study was funded by Dectrapharm, manufacturer of the RADT (Rapid Antigen Detection Test). Dectrapharm had no role in the design and conduct of the study, the collection, analysis, and interpretation of the data, or the review or approval of the manuscript. There are no patents, products in development or marketed products to declare. This does not alter the authors’ adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: jeremie.cohen@inserm.fr
clinical spectrum on sensitivity, specificity, likelihood ratios and predictive values of RADTs at the same time [13–15]. As well, RADT sensitivity is affected by inoculum size [16–18] – the amount of GAS colonies identified on throat culture and considered a proxy of the bacterial load on the swab. In vitro, the threshold of positivity of RADTs is between 10^5 and 10^7 colony-forming units per ml [19]. To date, several studies evaluated the effects of clinical spectrum and inoculum size on RADT sensitivity, but these effects were studied only separately, and no study has analysed the potential relation of clinical spectrum and inoculum size.

Because of the need to reduce antibiotics consumption, clinical decision rules were developed to help clinicians determine which patients should undergo testing and/or treatment with antibiotics. A decision rule based on the McIsaac score was recently proposed for adults and children [20–22], but this McIsaac-score–strategy was insufficiently validated in children [21–23]. In one validation study by McIsaac et al., the reader cannot evaluate the risk of selection bias because the distribution of scores in children was not reported [21]. Another validation study did not include children with low-risk scores (score ≤1) [22]. A third study aimed to validate the McIsaac score in children, but the rule was modified for application in low-resource settings and no longer integrated the use of a throat culture or RADT [23].

We aimed to determine whether the diagnostic accuracy of a RADT for GAS pharyngitis is affected by the clinical spectrum effect and/or bacterial inoculum size and to validate a McIsaac-score–based decision rule.

Methods

Study Design

This is a secondary analysis of data from an office-based, multicenter, prospective study that took place in France between October 2009 and June 2011 (unpublished data). The aims of the princeps study were to evaluate the frequency of GAS carriage in healthy children and to compare the performance of a RADT between children with pharyngitis and healthy children, with throat culture in a microbiology laboratory as the reference test. The intermediate results presented at the 28th Annual Meeting of The European Society for Paediatric Infectious Diseases, Nice, France, 2010; abstract A-229-0021-00920). This ancillary analysis was restricted to patients with pharyngitis from the first year of inclusion. The STARD statement was followed for reporting the results of the study [24].

Patients

Eligible patients were 3 to 15 years old who were evaluated by their paediatrician between October 2009 and May 2010, who had a diagnosis of pharyngitis and did not receive antibiotics for 7 days before inclusion. Pharyngitis was defined by inflammation of the pharynx and/or tonsils. In total, 17 French paediatricians who are part of a research and teaching network (ACTIV) participated in the study [25].

Throat Swabs

Throat samples were obtained by use of a double–swab collection–transportation system (Venturi Transystem Amies agar, COPAN Diagnostics, Corona, CA, USA). The RADT (StreptAt-est, Dectrapharm, France), a GAS-specific immuno-chromatographic strip assay, was performed immediately with swab #1 collected in the paediatrician’s office. Swab #2 was held at ambient temperature and sent within 72 hr to the Robert Debré Hospital laboratory by an express messenger service. On receipt, the swab was rolled over one-quarter of a trypticase soy agar plate with 5% sheep blood, and the inoculum was further distributed on the plate by streaking with a sterile wire loop. The plates were incubated anaerobically at 37°C and read after 18 to 24 hr. Negative cultures were reexamined after a additional 24 hr of incubation (48 hr total). ß-haemolytic colonies were further investigated by latex agglutination (Prolex, Pro-Lab Diagnostics, Richmond Hill, ON, Canada). With GAS positivity, inoculum size was estimated as follows: 1+, <10 colonies; 2+, 11–50 colonies; and 3+, >50 colonies per plate. Data for patients with 1+ and 2+ cultures were combined to produce a dichotomized variable (light or heavy GAS growth). Microbiologists were blinded to individual clinical data and RADT results.

Clinical Data

We collected clinical data needed to calculate the McIsaac score for each patient. The McIsaac score predicts GAS pharyngitis and ranges from 0 to 5 [20–22]. McIsaac criteria include age 3 to 15 years, history of fever, tonsillar swelling or exudate, tender anterior cervical adenopathy, and absence of cough. As for other studies [13,26], the adenopathy criterion was modified to include any anterior cervical adenopathy ≥1 cm and/or presence of tenderness because tender lymph nodes can be difficult to assess in children. As for other authors [13–15], extreme McIsaac scores were combined (scores 1 and 2, and 4 and 5), and for analysis of clinical components of the score, age was re-coded into 2 categories of equal range (3–8 years; 9–14 years).

Sample Size

In the original study, sample size was estimated so that the 95% confidence interval (95CI) for RADT sensitivity would be a +5 % estimation. Assuming sensitivity to vary between 75% and 95% and a GAS prevalence of 35%, a sample of 714 children with pharyngitis was needed. No other sample size calculation was performed for this secondary analysis.

Ethics

The study protocol was approved by the Institutional Review Board Comité de Protection des Personnes Ile-de-France XI (n°09016) and the French administrative authorities (CNIL, n°1354254; Afsaps, n°2009-A00066-51). Parent and patient approval for participation was obtained before inclusion. Data were double entered into 4D software version 6.4 (4D) and the database was fully anonymized.

Statistical Analysis

Hospital laboratory throat culture was considered the reference test. First, we described patient demographic features and overall prevalence of groups A, C and G ß-haemolytic streptococci. Patients with group C or G ß-haemolytic streptococci were considered GAS-negative in the analysis. Then, GAS prevalence by McIsaac score and each McIsaac criterion were described. Second, we used chi-square tests to investigate the clinical spectrum effect on RADT diagnostic accuracy (sensitivity, specificity, likelihood ratios and predictive values) by McIsaac score and each McIsaac criterion. Third, chi-square tests were used to compare the distribution of GAS inoculum size by office-to-laboratory delay (≤48 vs. >48 hr) and to study the effect of inoculum size (light or heavy GAS growth) on RADT sensitivity. Fourth, the distribution of heavy inocula by McIsaac score was explored by a chi-square test for trend, and the association of clinical signs and inoculum size was evaluated by comparing the frequency of heavy inoculum according to each clinical criterion.
by chi-square tests. Fifth, a binomial model with an identity link was used to evaluate the combination of clinical spectrum effect and inoculum size on RADT sensitivity [27]. Differences in RADT sensitivity were estimated by restricting the analysis to patients with a positive throat culture and using the result of the RADT as the outcome in the model. RADT sensitivity was investigated by univariate modeling according to each McIsaac criterion and then, multivariate modeling by all McIsaac criteria and inoculum size. Potential interaction between inoculum size and each McIsaac criterion was assessed by use of a Wald test. For multivariate modeling, variables were added one by one following the current univariate statistical significance order until non-convergence of the model. Significant interactions were managed by introducing corresponding interaction terms in the multivariate model and, in case of non-convergence, by stratification on inoculum size. Sixth, the accuracy of the following McIsaac-score–based decision rule was studied in terms of sensitivity and specificity: (a) scores <2, no further testing or antibiotic; (b) scores 2–3, culture all, antibiotics only with positive culture; (c) scores ≥4, treat empirically with antibiotics [22]. Statistical analysis involved use of Stata/SE 12 (StataCorp, College Station, TX, USA).

Results

Population Characteristics, GAS Prevalence

In total, 807 children met the inclusion criteria; 1 patient with unreadable RADT results, 2 with lost throat swabs and 19 with missing data for McIsaac score calculation were secondarily excluded. Therefore, analysis involved 785 children (351 girls [44.7%]), with mean (SD) age 6.1 (2.5) years. Overall, GAS prevalence was 36.3% (95CI, 32.9%–39.8%). In all, 4 (<1%) Group C and 7 (<1%) Group G β-haemolytic streptococci were identified on BAP. The mean age of secondarily excluded children was 5.9 (3.0) years and GAS was found in throat cultures for 4/20 (20.0% [7.9%–39.2%]).

GAS prevalence increased with increasing McIsaac score (p<0.001) (Tables 1 and 2), from 24.3% to 41.2% for children with ≤2 and ≥4 McIsaac score, respectively. Some clinical criteria were also associated with GAS: age <9 years, absence of cough, and presence of fever (Table 2).

### Table 1. Distribution of the McIsaac scores and corresponding frequencies of positive results of rapid antigen detection test (RADT), group A streptococcus (GAS) and heavy GAS inocula (N = 785).

<table>
<thead>
<tr>
<th>McIsaac score</th>
<th>Total</th>
<th>RADT-Positive</th>
<th>GAS-Positive</th>
<th>Heavy Inoculum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12 (1)</td>
<td>2/12 (17)</td>
<td>3/12 (25)</td>
<td>2/3 (67)</td>
</tr>
<tr>
<td>2</td>
<td>103 (13)</td>
<td>21/103 (20)</td>
<td>25/103 (24)</td>
<td>20/25 (80)</td>
</tr>
<tr>
<td>3</td>
<td>262 (33)</td>
<td>82/262 (31)</td>
<td>89/262 (34)</td>
<td>73/89 (82)</td>
</tr>
<tr>
<td>4</td>
<td>286 (36)</td>
<td>115/286 (40)</td>
<td>114/286 (40)</td>
<td>100/114 (88)</td>
</tr>
<tr>
<td>5</td>
<td>122 (16)</td>
<td>54/122 (44)</td>
<td>54/122 (44)</td>
<td>48/54 (89)</td>
</tr>
</tbody>
</table>

*All scores were ≥1 because all included patients were 3 to 15 years old.

bLaboratory throat culture positive for GAS.

cp = 0.09, χ² test for trend for the frequency of heavy inoculum by McIsaac score.

doi:10.1371/journal.pone.0039085.t001

Spectrum Effect

Overall, the sensitivity and specificity of the RADT was 86.7% (82.2%–90.4%) and 94.6% (92.2%–96.4%), respectively; positive and negative likelihood ratios 16.1 (11.1–23.2) and 0.14 (0.10–0.19), respectively; and positive and negative predictive values 90.1% (86.0%–93.4%) and 92.6% (89.9%–94.7%), respectively.

We found significant variations in RADT sensitivity and specificity by McIsaac score (Table 2). Sensitivity increased from 75% (55.1%–89.3%) to 84.3% (73.0%–91.1%) and to 89.9% (84.3%–94.0%; p = 0.02), and specificity decreased from 97.7% (91.9%–99.7%) to 96.0% (91.8%–98.4%) and to 92.5% (88.4%–95.5%; p = 0.04) with McIsaac scores ≤2, 3 and ≥4, respectively (Table 2). RADT likelihood ratios and predictive values varied, but not significantly, by McIsaac score.

The RADT performance varied significantly by McIsaac criteria. For example, sensitivity was significantly lower in children ≤9 years old than in younger children (66.7% [47.2%–82.7%] vs. 89.0% [84.5%–92.6%]; p<0.01). The negative predictive value varied, but not significantly, by McIsaac score or clinical criteria, ranging from 88.8% (80.8%–94.3%) for children without tonsillar swelling or exudate to 94.6% (90.0%–97.5%) for children without fever.

Effect of Inoculum Size on RADT Sensitivity

Among 285 GAS-positive throat cultures, 243 (85% [80.6%–89.2%]) showed heavy inoculum. The distribution of inoculum size did not differ by office-to-laboratory delay (≥48 vs. >48 hr, p = 0.97). RADT sensitivity varied widely and significantly between patients with light and heavy inoculum (40.5% [25.6%–56.7%] vs. 94.7% [91.0%–97.1%]; p<0.001). Light inoculum was found in 65.8% (50.0%–81.6%) of false-negative RADT cases versus 6.9% (3.7%–10.1%) of true-positive cases (p<0.001).

Frequency of heavy inocula increased with increasing McIsaac score, although not significantly (p = 0.09, table 1). Heavy inoculum was significantly more frequent in children <9 years old than in older children (87.1% [82.9%–91.2%] vs. 70.0% [52.6%–87.4%]; p<0.05).

Combined Effect of Clinical Spectrum and Inoculum Size

The results of univariate binomial modeling for RADT sensitivity were almost identical to those of stratified analysis (Tables 2 and 3): age ≥9 years and presence of cough were significantly associated with decreased RADT sensitivity (p<0.05). After stratification on inoculum size, wide and statistically significant variations in RADT sensitivity were found in patients with light inoculum (e.g., 37.4% variation in sensitivity by age, p<0.01), but not widely (<6%) nor significantly in patients with heavy inoculum. Interactions were found between inoculum size and age (p = 0.03), anterior cervical adenopathy (p = 0.01), and tonsillar swelling or exudate (p = 0.07).

Because of these multiple interactions and because of no convergence in the multivariate model after introducing interaction terms, 2 multivariate models of RADT sensitivity were fit by stratifying on inoculum size (Table 3). Results of multivariate modeling were close to those of univariate modeling. Light inoculum was associated with wide and statistically significant adjusted variation in RADT sensitivity (e.g., from 23.6% to 40.9%), but heavy inoculum was not widely (≤6%) or significantly associated.
Table 2. Spectrum effect analysis: performance of rapid antigen detection test by clinical spectrum.

<table>
<thead>
<tr>
<th>Midsaac score</th>
<th>n (%)</th>
<th>GAS (%)</th>
<th>Sensitivity (95CI)</th>
<th>Specificity (95CI)</th>
<th>p&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PLR (95CI)</th>
<th>p&lt;sup&gt;b&lt;/sup&gt;</th>
<th>NLR (95CI)</th>
<th>p&lt;sup&gt;b&lt;/sup&gt;</th>
<th>PPV (95CI)</th>
<th>p&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NPV (95CI)</th>
<th>p&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤2</td>
<td>115 (14)</td>
<td>24.3</td>
<td>&lt;0.001&lt;sup&gt;c&lt;/sup&gt;</td>
<td>75.0 (55.1–89.3)</td>
<td>0.02&lt;sup&gt;c&lt;/sup&gt;</td>
<td>97.7 (91.9–99.7)</td>
<td>0.04&lt;sup&gt;c&lt;/sup&gt;</td>
<td>32.6 (8.2–130.5)</td>
<td>0.23</td>
<td>0.26 (0.13–0.49)</td>
<td>0.09</td>
<td>91.3 (72.0–98.9)</td>
<td>0.61&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>262 (33)</td>
<td>34.0</td>
<td>84.3 (75.0–91.1)</td>
<td>20.8 (1.0–43.3)</td>
<td>0.16 (0.10–0.27)</td>
<td>91.5 (83.2–96.5)</td>
<td>0.92 (87.3–95.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥4</td>
<td>408 (51)</td>
<td>41.2</td>
<td>89.9 (843–940)</td>
<td>12.0 (7.7–18.7)</td>
<td>0.11 (0.07–0.17)</td>
<td>89.3 (83.7–93.6)</td>
<td>0.92 (88.9–95.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–8</td>
<td>663 (84)</td>
<td>38.5</td>
<td>&lt;0.01</td>
<td>89.0 (845–926)</td>
<td>0.02&lt;sup&gt;c&lt;/sup&gt;</td>
<td>94.1 (91.4–96.2)</td>
<td>0.45</td>
<td>15.1 (10.2–22.4)</td>
<td>0.62</td>
<td>0.12 (0.08–0.17)</td>
<td>&lt;0.001</td>
<td>90.4 (86.1–93.8)</td>
<td>0.48</td>
</tr>
<tr>
<td>9–14</td>
<td>122 (16)</td>
<td>24.6</td>
<td>66.7 (47.2–82.7)</td>
<td>20.4 (6.5–64.0)</td>
<td>0.34 (0.21–0.57)</td>
<td>87.0 (66.4–97.2)</td>
<td>0.89 (82.2–95.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of cough</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>428 (55)</td>
<td>44.4</td>
<td>&lt;0.001&lt;sup&gt;c&lt;/sup&gt;</td>
<td>90.0 (848–939)</td>
<td>0.02&lt;sup&gt;c&lt;/sup&gt;</td>
<td>94.1 (90.3–96.7)</td>
<td>0.65</td>
<td>15.3 (9.2–25.5)</td>
<td>0.89</td>
<td>0.11 (0.07–0.16)</td>
<td>0.02</td>
<td>92.4 (87.6–95.8)</td>
<td>0.07</td>
</tr>
<tr>
<td>No</td>
<td>357 (45)</td>
<td>26.6</td>
<td>80.0 (705–87.5)</td>
<td>16.1 (9.4–27.7)</td>
<td>0.21 (0.14–0.31)</td>
<td>85.4 (76.3–92.0)</td>
<td>0.92 (89.2–95.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior cervical adenopathy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>407 (52)</td>
<td>34.6</td>
<td>0.32</td>
<td>83.0 (757–888)</td>
<td>0.07</td>
<td>97.7 (93.2–99.2)</td>
<td>0.001</td>
<td>36.8 (16.6–81.4)</td>
<td>0.003</td>
<td>0.17 (0.12–0.25)</td>
<td>0.12</td>
<td>95.1 (89.7–98.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>Yes</td>
<td>378 (48)</td>
<td>38.1</td>
<td>90.3 (842–946)</td>
<td>10.1 (6.7–15.2)</td>
<td>0.11 (0.06–0.18)</td>
<td>86.1 (79.5–91.2)</td>
<td>0.93 (89.9–96.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tonsillar swelling or exudates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>157 (20)</td>
<td>42.0</td>
<td>0.10</td>
<td>83.3 (721–914)</td>
<td>0.36</td>
<td>95.6 (91.1–98.8)</td>
<td>0.80</td>
<td>19.0 (7.2–49.7)</td>
<td>0.71</td>
<td>0.17 (0.10–0.30)</td>
<td>0.38</td>
<td>93.2 (83.5–98.1)</td>
<td>0.47</td>
</tr>
<tr>
<td>Yes</td>
<td>628 (80)</td>
<td>34.9</td>
<td>87.7 (826–917)</td>
<td>15.6 (10.5–23.3)</td>
<td>0.13 (0.09–0.19)</td>
<td>89.3 (84.4–93.1)</td>
<td>0.93 (90.6–95.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fever</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>246 (31)</td>
<td>31.3</td>
<td>0.049</td>
<td>88.3 (790–945)</td>
<td>0.62</td>
<td>93.5 (88.7–96.7)</td>
<td>0.43</td>
<td>13.6 (7.6–24.2)</td>
<td>0.47</td>
<td>0.13 (0.07–0.23)</td>
<td>0.66</td>
<td>86.1 (76.5–92.8)</td>
<td>0.15</td>
</tr>
<tr>
<td>Yes</td>
<td>539 (69)</td>
<td>38.6</td>
<td>86.1 (806–905)</td>
<td>17.8 (11.0–28.8)</td>
<td>0.15 (0.10–0.21)</td>
<td>91.8 (870–95.2)</td>
<td>0.91 (88.1–94.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>785</td>
<td>36.3</td>
<td>86.7 (822–904)</td>
<td>16.1 (11.1–23.2)</td>
<td>0.14 (0.10–0.19)</td>
<td>90.1 (860–93.4)</td>
<td>0.92 (89.9–94.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: GAS, group A streptococcus; PLR, positive likelihood ratio; NLR, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

<sup>a</sup>χ² test; <sup>b</sup>Mantel-Haenszel's χ² test; <sup>c</sup>χ² test for trend.

doi:10.1371/journal.pone.0039085.t002
Performance of the McIsaac-score–based Decision Rule

Among 785 children, 522 (66%) would have received antibiotics according to the decision rule. The rule had 98.9% (97.0%–99.8%) sensitivity and 52.0% (47.5%–56.5%) specificity.

Discussion

Accuracy of a diagnostic test can vary across patient subgroups within a population, a phenomenon referred to as spectrum effect (or spectrum bias) [10–12]. In assessing the relation of clinical spectrum and inoculum size with RADT performance for GAS (or spectrum bias) [10–12], we cannot conclude on whether these patients with low McIsaac scores and light inocula were more likely to be GAS carriers rather than truly infected patients.

We chose to investigate the McIsaac score in 3 categories rather than compare the accuracy of the RADT above and below a defined breakpoint [15] because this corresponds to the original aims of the score – to stratify patients into 3 levels of risk and to suggest a subsequent course of action (low-risk: no test, no antibiotics; intermediate risk: culture or RADT, antibiotics only if gas positive; high-risk: empirical antibiotic treatment without testing) [20–22]. According to our results, the McIsaac score alone does not allow for ruling out or affirming the diagnosis of GAS pharyngitis because the reported GAS prevalence ranged from 25% (too-high level to rule out) to 44% (too-low level to affirm) in children with a score of 1 and 5, respectively. Moreover, the McIsaac-score–based decision rule had only 52.0% (47.5%–56.5%) specificity, which does not seem consistent with the current need to reduce antibiotic consumption to contain antimicrobial resistance.

One of the strengths of this study is its prospective, multicenter design, which limits the selection bias potentially present in other studies [13,14]. All included children underwent RADT and laboratory culture, which excluded the possible indication and (greater inoculum size) might be associated with more intense symptoms (higher McIsaac score) (p = 0.09), although we might not have had enough statistical power to validate this hypothesis. However, because we were not able to differentiate GAS carriers from truly infected patients (i.e., by assessment of streptococcal antibody response) [16,28], we cannot conclude on whether these patients with low McIsaac scores and light inocula were more likely to be GAS carriers rather than truly infected patients.

Table 3. Analysis of combined spectrum and inoculum size effect on sensitivity of rapid antigen detection test (RADT).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate analysis*</th>
<th>Multivariate analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All inocula (n = 285)</td>
<td>Light inoculum (n = 42)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–8</td>
<td>ref</td>
<td>ref</td>
</tr>
<tr>
<td>9–14</td>
<td>–22.4 (−39.7; −5.1)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Lack of cough</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>ref</td>
<td>ref</td>
</tr>
<tr>
<td>No</td>
<td>–10.0 (−19.1; −9.0)</td>
<td>0.09</td>
</tr>
<tr>
<td>Anterior cervical adenopathy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>ref</td>
<td>ref</td>
</tr>
<tr>
<td>Yes</td>
<td>7.3 (−0.6; 15.2)</td>
<td>0.07</td>
</tr>
<tr>
<td>Tonsillar swelling or exudate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>ref</td>
<td>ref</td>
</tr>
<tr>
<td>Yes</td>
<td>4.3 (−5.6; 14.3)</td>
<td>0.40</td>
</tr>
<tr>
<td>Fever</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>ref</td>
<td>ref</td>
</tr>
<tr>
<td>Yes</td>
<td>–2.3 (−10.8; 6.3)</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Data RADT sensitivity difference (95% confidence interval) unless indicated.

Abbreviations: ref, reference; n/a, not applicable.

*Interaction test between clinical variables and inoculum size (Wald test).

Performance of the McIsaac-score–based Decision Rule

Among 785 children, 522 (66%) would have received antibiotics according to the decision rule. The rule had 98.9% (97.0%–99.8%) sensitivity and 52.0% (47.5%–56.5%) specificity.

Discussion

Accuracy of a diagnostic test can vary across patient subgroups within a population, a phenomenon referred to as spectrum effect (or spectrum bias) [10–12]. In assessing the relation of clinical spectrum and inoculum size with RADT performance for GAS pharyngitis, we confirm that the spectrum effect has an impact on RADT sensitivity and report for the first time that it also affects the predictive value of the RADT, which remained stable at about 92%, regardless of McIsaac score or its components. Clinicians do not want to expose their patients to GAS suppurative and therefore need a diagnostic test with high, stable negative predictive value. In this population with a GAS prevalence of 99.8% sensitivity and 52.0% specificity.

Performance of the McIsaac-score–based Decision Rule

Among 785 children, 522 (66%) would have received antibiotics according to the decision rule. The rule had 98.9% (97.0%–99.8%) sensitivity and 52.0% (47.5%–56.5%) specificity.

Discussion

Accuracy of a diagnostic test can vary across patient subgroups within a population, a phenomenon referred to as spectrum effect (or spectrum bias) [10–12]. In assessing the relation of clinical spectrum and inoculum size with RADT performance for GAS pharyngitis, we confirm that the spectrum effect has an impact on RADT sensitivity and report for the first time that it also affects the predictive value of the RADT, which remained stable at about 92%, regardless of McIsaac score or its components. Clinicians do not want to expose their patients to GAS suppurative and therefore need a diagnostic test with high, stable negative predictive value. In this population with a GAS prevalence of 99.8% sensitivity and 52.0% (47.5%–56.5%) specificity, which does not seem consistent with the current need to reduce antibiotic consumption to contain antimicrobial resistance.

One of the strengths of this study is its prospective, multicenter design, which limits the selection bias potentially present in other studies [13,14]. All included children underwent RADT and laboratory culture, which excluded the possible indication and
that the target sensitivity of diagnostic strategies for GAS mon in North America and Western Europe [29], we advocate remain subject to personal interpretation. Because GAS should be tested are not integrated in an explicit decision rule and tested for GAS infection manifestations highly suggestive of viral infection […] generally should not be

Another limitation to our study is that some throat swabs were plated >48 hr after collection. This delay had no significant effect on the distribution of GAS inocula, but prolonged or inadequate shipping conditions could have resulted in the loss of viability of GAS if the original swab had only small numbers of bacteria. Therefore, some swabs with light inocula could have been falsely read as negative on throat culture. Because RADTs are more likely to give negative results with light inocula, these swabs might have led to a systematic decrease in number of RADT false-negative results and systematic increase in number of RADT true-negative results, with over-estimation of the negative predictive value of the RADT as a result. We can assume that this bias occurred at random because shipping and all bacteriologic investigations were blinded to clinical data and RADT results – non-differential measurement bias. Such a bias usually leads to loss of power, and our results regarding the stability of the negative predictive value of the RADT should be considered with caution until they are confirmed with a larger sample of patients.

The American Academy of Pediatrics advises that “Children with manifestations highly suggestive of viral infection […] generally should not be tested for GAS infection” [7]. Such guidelines cannot be formally validated because the criteria proposed to determine patients who should be tested are not integrated in an explicit decision rule and remain subject to personal interpretation. Because of GAS carriage and because GAS complications have become uncommon in North America and Western Europe [29], we advocate that the target sensitivity of diagnostic strategies for GAS pharyngitis should not be 100%. Moreover, discussions should not focus on sensitivity alone. We showed that the negative predictive value of a test can be stable across patient subgroups because variations in prevalence can be balanced by concomitant variations in sensitivity and specificity. Therefore, clinicians who want to rule out GAS can rely on negative RADT results regardless of clinical features if they accept that about 10% of children with negative RADT results will have a positive throat culture. However, these results should be considered with caution because they depend highly on the study population, the training of clinical and laboratory personnel and the microbiological devices and protocols used. Moreover, available data suggest that false-negative RADT results are as likely to occur in truly infected patients as in GAS carriers [16,20]. Therefore, such a policy would be more acceptable in populations with very low incidence of supplicative and non-suppurative complications of GAS infection [30]. The McIsaac-score–based decision rule is insufficiently accurate for children, and efforts are needed to develop more specific selective testing strategies for GAS pharyngitis in children.

References


