Positive Predictive Value of Gen-Probe APTIMA Combo 2 Testing for Neisseria gonorrhoeae in a Population of Women with Low Prevalence of N. gonorrhoeae Infection

Matthew R. Golden,1,3 James P. Hughes,2 Linda E. Cles,1 Karen Crouse,4 Katherine Gudgel,1 Jinxin Hu,1 Paul D. Swenson,1 Walter E Stamm,1 and H. Hunter Handsfield2,3

1Division of Infectious Diseases and 2Department of Biostatistics, University of Washington, and 3Public Health–Seattle and King County, Seattle, 4Spokane Regional Health District Laboratory, Spokane, and 5Washington State Department of Health, Olympia, Washington

We collected specimens from women who tested positive for Neisseria gonorrhoeae using the Gen-Probe APTIMA Combo 2 (AC2), and we performed confirmatory tests using a nucleic acid amplification test (NAAT) that targets alternate gonococcal nucleic acid sequences. Among 59,664 specimens, 280 (0.47%) had positive results using AC2; 265 of these specimens were tested using the confirmatory NAAT, of which 258 yielded positive results (positive predictive value, 97.4%; 95% confidence interval, 95.1%–98.8%). Routine confirmatory testing of specimens with positive AC2 gonorrhea test results is not indicated.

Programs to test asymptomatic, sexually active young women for genital tract infections with Neisseria gonorrhoeae and Chlamydia trachomatis are widespread in many industrialized countries. In most settings, gonorrhea is much less common than chlamydial infection; 5.6% of female family planning clinic patients aged 14–24 years in the United States tested positive for C. trachomatis in 2002, whereas the prevalence of N. gonorrhoeae in that population was <1% in 2000 [1, 2].

Nucleic acid amplification tests (NAATs) are commonly used to test for C. trachomatis and gonorrhea [3, 4]. The Gen-Probe APTIMA Combo 2 (AC2) assay detects both C. trachomatis and N. gonorrhoeae. The AC2 package insert states that the test for N. gonorrhoeae has a sensitivity of 99.2% and a specificity of 98.7% when cervical swab specimens are used and a sensitivity of 91.3% and a specificity of 99.3% when urine specimens obtained from women are used. These estimates are comparable to those for other NAATs [5] and, if correct, would mean that the positive predictive values (PPVs) for the test in a population with a prevalence of N. gonorrhoeae of 1% would be 43% and 57% for cervical and urine specimens, respectively. Most clinicians would not consider this PPV adequate for a test result to be regarded as diagnostic, and the Centers for Disease Control and Prevention (CDC) recommends that confirmatory testing “be considered” when nonculture tests have a PPV of <90% [6]. However, estimates of the sensitivity and specificity of NAAT are not sufficiently precise to confidently define the PPV of these tests for N. gonorrhoeae in the very low-prevalence populations in which they are typically used.

To assess whether there is a need to routinely perform confirmatory tests for specimens that test positive for N. gonorrhoeae using AC2, we used the commercially available Gen-Probe APTIMA GC (AGC) assay to retest all specimens obtained from women that tested positive for N. gonorrhoeae by AC2 in a low-prevalence population. Because the study design we employed is not typically used to assess new diagnostic tests, we also calculated the number of specimens we would have had to test using both assays to obtain a similar estimate of the PPV had we enrolled patients in a study designed to precisely define the AC2’s sensitivity and specificity and then used those parameters to calculate the PPV.

Methods. Study specimens included consecutive urine and cervical swab specimens obtained from women who tested positive for N. gonorrhoeae in 3 public health laboratories and 1 academic laboratory in Washington state in 2002 and 2003. Specimens from sexually transmitted disease clinics were excluded. All specimens were collected as part of routine clinical care.

AC2 tests were initially performed on pooled lots of 4 or 5 specimens. AC2 testing was done according to the manufacturer’s instructions for individual specimen testing. If a pool yielded positive results, specimens were tested individually. Some pools contained both urine and cervical swab specimens. Only specimens that tested positive individually were regarded as positive; these specimens were frozen at −20°C for up to 30 days and at −70°C thereafter for future evaluation.

AGC was used to test all specimens that initially yielded positive AC2 results, as well as to test a sample of consecutive individual specimens at each laboratory that originally tested
negative for *N. gonorrhoeae*. AGC is similar to AC2 except that it uses oligonucleotides that target 16S rRNA sequences different from those of AC2.

Ninety-five percent confidence intervals for the estimated PPV were calculated using an exact binomial method [7]. The sample size for the confirmatory testing design in table 1 was determined by finding the smallest N that consistently (i.e., 95% of the time) gave a lower bound of a 95% CI for PPV of \( \geq 0.90 \). Technically,

\[
N_{\text{confirm}} = \min_n \sum I(0.9 < CI(\alpha,v,n,x)) \left( \binom{n}{x} v^x (1 - v)^{n-x} \geq 0.95 \right),
\]

where \( I(B) \) is the indicator function with value 1 if B is true and 0 otherwise, and \( CI(\alpha,v,n,x) \) is an \( \alpha \) level (exact) confidence interval for the PPV, \( v \), based on a sample of size \( n \) with \( x \) true positives. \( N_{\text{dual}} \) and \( N_{\text{confirm}} \) were related by

\[
N_{\text{dual}} = \frac{N_{\text{confirm}}}{p \times \text{sensitivity} + (1 - p) \times (1 - \text{specificity})},
\]

where \( p \) is the disease prevalence. PPV and specificity were related using

\[
\text{PPV} = \left( 1 + \frac{1 - \text{specificity}}{\text{sensitivity}} \right) \left( \frac{1 - p}{p} \right)^{-1}.
\]

Investigators obtained a certificate of exemption from the University of Washington Institutional Review Board (Seattle) for the study.

**Results.** The 4 laboratories tested 59,664 unduplicated cervical or urine specimens during the study period, of which 280 (0.47%) tested positive for *N. gonorrhoeae* (table 2). Of these 280 positive specimens, 265 (95%) were available for confirmatory testing, of which 258 had positive results of the AGC test. The PPV of AC2 for *N. gonorrhoeae* was 97.4% (95% CI, 95.1%–98.8%). All specimen pools that yielded positive results included \( \geq 1 \) specimen that yielded positive results when tested individually. All 194 specimens that tested negative for gonorrhea using AC2 had negative AGC results.

Assuming a true PPV of 97% and a prevalence of 0.5%, a total of 33,195 specimens would have had to be tested using both AC2 and AGC to define AC2’s sensitivity and specificity with sufficient precision to demonstrate that the lower bound of the 95% CI for the test’s PPV was \( >90\% \). In a population with a 5% prevalence of *N. gonorrhoeae* infection, 3319 specimens would need to be tested using both tests (table 2). In contrast, only 154 AC2-positive specimens would need to be tested using AGC to similarly define AC2’s PPV in a study limited to confirmatory testing, the approach we used here.

**Discussion.** Most tests for *N. gonorrhoeae* performed in the United States are now done using technologies other than culture, and many are performed using NAATs [3]. These assays are typically used in populations with a very low prevalence of *N. gonorrhoeae* infection [2]. However, the PPVs of nonculture tests are not defined in such populations, and false-positive results have been reported using PCR [8].

As demonstrated by our analysis, conventional studies (i.e., those in which all specimens are assessed using 2 different tests) cannot practically define the specificity of nonculture tests with sufficient precision to confidently estimate their PPV in a population with a very low prevalence of infection. The CDC advocates that confirmatory tests be considered when nonculture tests are used in such populations. An alternative to universal confirmatory testing is to directly estimate the PPV in a very low-prevalence population by performing confirmatory tests

---

**Table 1. Number of specimens needed in a confirmatory testing design and dual testing design to have 95% probability that the lower bound of a 95% CI for positive predictive value (PPV) is \( \geq 0.90 \).**

<table>
<thead>
<tr>
<th>True PPV, %</th>
<th>( N_{\text{confirm}} )</th>
<th>( N_{\text{dual}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>93</td>
<td>1108</td>
<td>228,987</td>
</tr>
<tr>
<td>94</td>
<td>583</td>
<td>121,782</td>
</tr>
<tr>
<td>95</td>
<td>352</td>
<td>74,311</td>
</tr>
<tr>
<td>96</td>
<td>231</td>
<td>49,280</td>
</tr>
<tr>
<td>97</td>
<td>154</td>
<td>33,195</td>
</tr>
<tr>
<td>98</td>
<td>100</td>
<td>21,778</td>
</tr>
<tr>
<td>99</td>
<td>70</td>
<td>15,400</td>
</tr>
</tbody>
</table>

**NOTE.** Sensitivity was assumed to be 90%. Dual testing design \( (N_{dual}) \) involves testing of all specimens using 2 different tests. Confirmatory testing \( (N_{\text{confirm}}) \) involves performing a second test only on specimens that test positive with an initial test.
Table 2. Outcomes of Neisseria gonorrhoeae confirmatory testing study among women tested outside of sexually transmitted disease clinics.

<table>
<thead>
<tr>
<th>Laboratory or specimen type</th>
<th>No. of specimens tested during study period</th>
<th>No. (%) of positive AC2 results</th>
<th>No. (%) of AC2-positive specimens tested using AGC and AGC test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All</td>
<td>Positive</td>
</tr>
<tr>
<td>Laboratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spokane Public Health</td>
<td>8483</td>
<td>19 (0.2)</td>
<td>19</td>
</tr>
<tr>
<td>UWa</td>
<td>13,962</td>
<td>75 (0.5)</td>
<td>40 (50)</td>
</tr>
<tr>
<td>PHSKCa</td>
<td>4024</td>
<td>48 (1.2)</td>
<td>68</td>
</tr>
<tr>
<td>WA State</td>
<td>33,195</td>
<td>138 (0.4)</td>
<td>138</td>
</tr>
<tr>
<td>All</td>
<td>59,664</td>
<td>280 (0.5)</td>
<td>265</td>
</tr>
<tr>
<td>Specimen type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swab</td>
<td>NA</td>
<td>NA</td>
<td>200</td>
</tr>
<tr>
<td>Urine</td>
<td>NA</td>
<td>NA</td>
<td>65</td>
</tr>
</tbody>
</table>

NOTE. AC2, APTIMA Combo 2 (Gen-Probe); NA, not available; AGC, APTIMA GC (Gen-Probe); PHSKC, Public Health Seattle & King County (Washington); UW, University of Washington (Seattle).

a Some specimens were tested at both the UW laboratory and the PHSKC laboratory. In these instances, confirmatory tests were performed only at the PHSKC laboratory. Twenty-one specimens originally tested at the UW laboratory had confirmatory testing done at the PHSKC laboratory.

b All specimens were also culture positive for N. gonorrhoeae.

on all specimens that yield positive results for a defined period of time. Results from such studies could guide a decision on whether routine confirmatory testing is warranted [9].

In this study, we employed such an approach and demonstrated that AC2 has a PPV of 97% in a population of women in which the prevalence of N. gonorrhoeae infection is 0.5%. Although future studies should seek to replicate our results, on the basis of our findings, we do not believe that use of the AC2 for detection of N. gonorrhoeae requires routine confirmatory testing.

Our study has 4 limitations. First, we studied a single NAAT in women. Our findings should not be extrapolated to indicate that confirmatory testing is not needed when using other NAATs, nor do our findings address the issue of AC2 performance in men or on vaginal swab specimens. Second, our laboratories initially tested pooled specimens, a standard procedure in the 4 study laboratories that is neither approved by the US Food and Drug Administration nor recommended by APTIMA’s manufacturer. Because we initially tested pooled specimens, all study specimens were tested twice before confirmatory testing was performed, a procedure that might affect test specificity and, consequently, PPV. However, none of our pooled specimens yielded positive results and then failed to yield a positive result for an individual specimen, a finding that might have raised concerns that repeated testing improved specificity. Increased false-positive testing could be more common when testing individual specimens, because the number of tests performed increases, or could increase with pooling due to increased specimen handling and opportunity for contamination. Thus, we cannot say with certainty how specimen pooling might have affected our estimate of PPV. Third, we assumed that specimens that yielded negative AGC results had false-positive AC2 results. If AGC is less sensitive than AC2, the true PPV of AC2 may be higher. Fourth, and most importantly, the approach we used would not have ruled out all false-positive results. False-positive NAAT results can occur in 3 ways. First, carry-over amplicon can contaminate samples from a previous reaction. Because the AGC assay does not amplify or detect AC2 amplicon, the confirmatory approach we used excludes false-positive results stemming from amplicon contamination. Second, if the rRNA sequences targeted by AC2 and AGC were both present in an organism other than N. gonorrhoeae, some specimens we classified as having confirmed positive results may have had false-positive results with both tests. The AC2 package insert [10] information and unpublished data suggest that both AC2 and AGC are highly specific (J. Schachter and D. V. Ferrero, personal communication). Although we cannot completely exclude the possibility that we underestimated the PPV for AC2 because of some shared non-specificity of AC2 and AGC, routine confirmatory testing using either AGC or another NAAT would not necessarily overcome this limitation. Third, false-positive test results can result from mislabeling of specimens or specimen contamination with organisms. Elimination of such false-positive results would require patients to provide additional specimens.

In summary, we have shown that the AC2 does not require routine confirmatory testing, even in populations with a very low prevalence of N. gonorrhoeae infection. However, clinicians should be aware that all tests can have false-positive results, and repeated or confirmatory testing may be indicated when
specific clinical circumstances lead clinicians to doubt laboratory results. The study design we used should be considered for wider application in evaluating new diagnostic tests for use in low-prevalence populations.

Acknowledgments

We thank Dr. Craig Hill for comments on this manuscript.

Financial support. National Institutes of Health (grant NIH K23 AI01846; to M.R.G.) and an unrestricted grant from Gen-Probe.

Potential conflicts of interest. M.R.G. and P.S. received recent research funding from Gen-Probe (submitted article only). All other authors: No conflict.

References