CDC reports on two alternative HIV testing algorithms

Amy Carpenter Aquino

August 2019—For HIV testing, a three-step algorithm that differs from the one recommended since 2014 can potentially reduce the number of tests performed and speed up the availability of viral load results, according to a CDC analysis presented at the HIV Diagnostics Conference in March. The CDC also evaluated a two-step algorithm that begins with the BioPlex 2200 HIV Ag-Ab differentiation assay and ends with the Aptima HIV-1 Quant assay.

The currently recommended CDC/APHL algorithm calls for three steps: HIV-1/2 antigen-antibody screening, followed by an HIV-1/2 antibody differentiation assay, and an HIV-1 nucleic acid test, if needed, for indeterminate or nonreactive results in the second step.

“At this point, the 2014 recommended algorithm has been implemented pretty widely in laboratories across the United States,” said Marc Pitasi, MPH, epidemiologist, Division of HIV/AIDS Prevention, CDC, in his presentation at the conference. Implementation challenges remain, however, mostly involving the use of a supplemental differentiation assay as step two and accessibility, resource requirements, and potential delays in turnaround time related to using qualitative NAT in step three. “Given those challenges, some laboratories might benefit from an alternative approach.”

One approach is to use a quantitative HIV nucleic acid test for HIV-1 viral load as step two, followed by antibody differentiation testing as step three if the viral load result is target not detected. (See “An alternative laboratory testing algorithm.”)

“Our objective was to provide data that could help demonstrate the potential feasibility of this alternative algorithm,” Pitasi said of the revised three-step algorithm. The CDC evaluated its performance against the 2014 algorithm using specimens that had been tested with five HIV-1/2 antigen-antibody combination screening assays and two HIV-1 viral load quantitative assays.

Specimens were collected from more than 6,000 patients of unknown HIV status but at high risk for HIV infection who were seeking testing at two clinics in Los Angeles between 2003 and 2005. Since specimen volume was limited, screening tests were performed on serum, and viral load tests were performed on plasma.

The five antigen-antibody screening tests used were the Abbott Architect HIV Ag/Ab Combo, Bio-Rad GS HIV Combo Ag/Ab EIA, Siemens Advia Centaur HIV Ag/Ab Combo, Bio-Rad BioPlex 2200 HIV Ag-Ab, and Alere Determine HIV-1/2 Ag/Ab Combo. All specimens with reactive results for any of the antigen-antibody screening tests were included in the analysis.

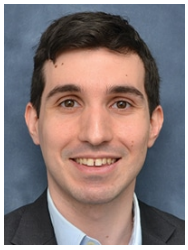
Specimens were followed through the alternative algorithm using a viral load quantitative test—Roche Amplicor HIV-1 Monitor or Hologic Aptima HIV-1 Quant—as step two. (Neither is FDA approved for diagnosis.) Specimens with no detected viral load were then tested with the Bio-Rad Geenius HIV 1/2 Supplemental Assay.

Specimens were classified according to the 2014 recommended algorithm. Five specimens showed early infection, defined as nonreactive/indeterminate Ab Geenius results and reactive Aptima qualitative assay results; 152 specimens showed established infection, defined as reactive Geenius results; and 38 specimens had false-positive screening results, defined as nonreactive/indeterminate Ab Geenius results and nonreactive Aptima qualitative assay results.

“We calculated the percentage of specimens that was correctly classified by antigen-antibody screening followed by viral load testing, relative to the recommended diagnostic algorithm individually for each screening and viral load test, and for any combination of these tests,” Pitasi said. Sensitivity for detecting early infection was between 53 and 60 percent for the Abbott, Siemens, and Bio-Rad BioPlex and Combo Ag/Ab assays, while the Alere assay showed a slightly lower sensitivity of 40 percent. “The confidence intervals were quite wide,” he noted.

For established infection, sensitivity and specificity were uniformly high among all five antigen-antibody screening tests.

Looking at viral load assay performance, the Roche Amplicor assay correctly classified all of the specimens with early infection detected by the antigen-antibody screening tests. In other words, they had quantified viral loads, Pitasi said. The Roche assay also correctly classified 95 percent of the 152 specimens with established infection. The proportion of detectable viral load was similar regardless of which screening test was used.



'All infected persons, of course, would go on to get a viral load eventually. We're just shortening that process.' Marc Pitasi, MPH

"Finally, all of the specimens [38] that were false reactive on any screening test were correctly classified, which in this case means they had a Roche viral load result of target not detected," he said.

The Aptima Quant also correctly classified all of the specimens with early infection and all 110 specimens with established infection. Because of limited specimen volume, false-positive specimens were not tested with the Aptima Quant.

Eight specimens with established infection had undetectable virus on the Roche viral load test. "If the alternative algorithm were used here, these would go on to test with Geenius as the third step," he said. There are no Aptima quantitative assay results on those eight specimens.

"Following the alternative algorithm, we saw that each of the eight specimens first was positive on every screening test, with the exception of the one specimen that was false-negative on the [Alere] Determine assay but positive on the other four screening tests."

In step two, Roche Amplicor gave results of target not detected, which was considered a false-negative for diagnostic purposes.

The eight specimens then proceeded to testing on the Geenius assay as step three and showed HIV-1 positive results. Since all specimens tested positive on the Geenius assay, "each of these specimens would have been detected using the alternative algorithm," with the possible exception of the specimen that tested false-negative on the Alere Determine assay. "With that same exception, they also would have all been detected using the recommended algorithm," he added.

In shifting the differentiation supplemental testing to step three, however, only those specimens would require differentiation supplemental tests, compared with 152 tests using the recommended algorithm, Pitasi said.

"Keep in mind that all infected persons, of course, would go on to get a viral load eventually. We're just shortening

that process."

Among the 38 specimens that had false reactive results on any of the five screening tests, eight specimens were false reactive on more than one screening test. "Using the recommended algorithm, all of these specimens would have proceeded to Geenius, followed by the Aptima qualitative assay, where they would have been resolved as negative," Pitasi said. "Using the alternative algorithm, these specimens would have proceeded first to a viral load test, which yielded the result of target not detected for all eight."

One specimen had a result of HIV-1 indeterminate on the Geenius assay and would likely require additional testing relative to the recommended algorithm. Cross-contamination or other sources of error could have explained that result, he suggested.

Of the remaining 30 of 38 specimens, four had Geenius assay results of HIV-2 indeterminate, which might require additional testing if the alternative algorithm were used. "Of these four specimens, three were false reactive on Abbott Architect, and one was false reactive on BioPlex," he said, noting that all four specimen results had relatively low signal-to-cutoff values.

Pitasi noted the study's limitations. Tests were performed on stored specimens, collected between 2003 and 2005; antigen-antibody screening and Ab supplemental testing were performed between 2008 and 2015 and viral load testing was performed in 2017. For the eight specimens not detected on viral load testing, "there could have been RNA degradation." All positive specimens in the sample had high viral loads (more than 1,000 copies), so there was no investigation of the sensitivity of the alternative algorithm in lower viral load specimens. Limited specimen volume meant there was no repeat testing of initial reactive results for most of the antigen-antibody screening tests, which could have increased the number of false reactive specimens. Last, the Geenius tests were performed using software version 1.1, which was updated in 2017 to raise the cutoff value of the HIV-2 gp140 band.

"In closing, this alternative algorithm performs well overall with the vast majority of early and established infection confirmed by viral load testing at the second step," Pitasi said.

Using the alternative algorithm would have averted antibody supplemental testing in more than 75 percent of the specimens examined, he said. "However, of the 46 specimens that did go on to Geenius, five of them had indeterminate results, including results of HIV-2 indeterminate, which might have required additional testing."

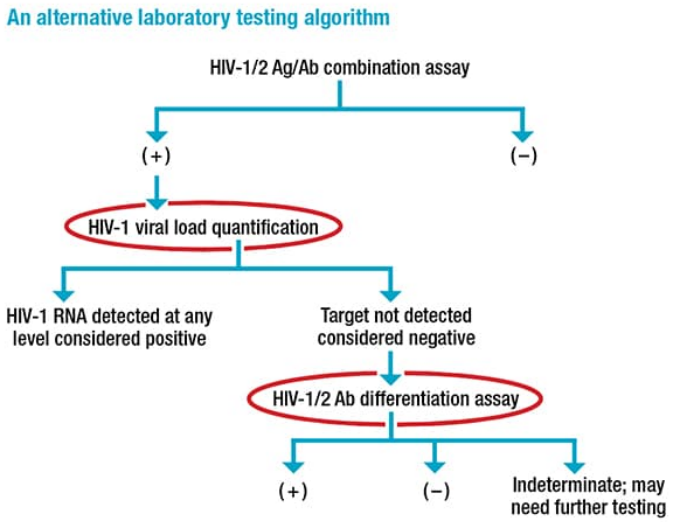
Although the alternative algorithm would likely reduce the overall number of specimens with difficult-to-interpret Geenius results, a few of those specimens would still remain and could potentially trigger additional testing or blood draws.

"This alternative algorithm has the potential to reduce the total number of tests performed, avoid the potentially lengthy turnaround time related to obtaining Aptima Qual results, especially in laboratories that don't do this in-house. It can also potentially expedite the availability of viral load results to improve patient care," Pitasi said.

However, there remains a need for better understanding of cost-effectiveness and feasibility and barriers to implementing this algorithm and to using viral load tests for diagnosis.

"Most importantly, we still need an FDA-approved quantitative viral load test that also has an indication for diagnosis of HIV-1 infection before an algorithm like this one could be implemented more widely."

Silvina Masciotra, MS, research microbiologist with the CDC's Division of HIV/AIDS Prevention, presented an alternative two-step algorithm that also used HIV-1 RNA as the second step. "HIV-2 infections are rare in the United States," she said in reference to the HIV-1/2 differentiation test, which is step two in the CDC/APHL recommended three-step algorithm.



Despite the widespread adoption of the recommended algorithm, “fewer labs have implemented the use of the FDA-approved nucleic acid test for the detection of acute infection,” for workload reasons, Masciotra said. “Viral load is often used as an alternative third test, though it is not an intended use.”

The CDC took part in the clinical trial for the BioPlex 2200 HIV Ag-Ab assay, a multiplex flow immunoassay intended for the simultaneous detection of p24 and HIV-1 and HIV-2 antibodies in human serum or plasma. “We had the chance to evaluate seroconversion panels at the CDC,” Masciotra noted. Those evaluations found that the BioPlex assay performed similarly to any other FDA-approved laboratory-based antigen-antibody immunoassay in detecting HIV-1 early infection.

The Hologic Aptima HIV-1 Quant assay on the Panther system is FDA approved for viral load monitoring. “We know the same assay is approved outside of the United States with a dual claim” for HIV-1 diagnosis and monitoring, Masciotra said. “This means we can use the viral load assay for diagnosis by following the interpretation” provided in the package insert that is approved outside of the United States. “Everything that was detected, even if it was not quantified because it was lower than the limit of quantification, we called it HIV-1 RNA positive,” Masciotra said.

The Aptima Quant assay is a high-throughput, fully automated, test tube platform with random access and uses transcription-mediated amplification. The reported limit of detection is 12 copies/mL. Its linear range of quantification is 30–107 copies/mL. The CDC did an in-house evaluation of the Aptima Quant assay for diagnosis.

In a comparison study of 417 samples from U.S. seroconverters infected with HIV-1, the Aptima Quant detected virus in more samples including the seronegative phase than the Hologic Aptima HIV-1 RNA Qualitative Assay approved for diagnosis. In patients with established HIV-1 infection, both tests performed similarly. “In our validation, we concluded that both assays were equal.”

Why use HIV-1 RNA viral load as the second test in the algorithm? “Since we do a lot of evaluations in the laboratory, I had the data. I said, ‘What if you look at a two-test diagnostic algorithm using a screening test that gives you the ability to differentiate HIV-1 from HIV-2 antibodies and p24, followed by the HIV-1 nucleic acid test—in this case the Aptima Quant—that also has approval outside of the United States,’” and may be approved by the FDA someday.

The objective was to compare the performance of a two-test diagnostic algorithm, consisting of screening with an antigen-antibody HIV-1/2 differentiation immunoassay, followed by an HIV-1 nucleic acid test, to the currently recommended three-test algorithm.

The analysis of specificity for the BioPlex assay during the clinical trial was calculated using 596 HIV-negative samples and the CDC contributed some of the testing. For the Aptima Quant assay, the laboratory tested 478

HIV-1–negative samples. “Because contamination problems have been reported with viral load open platforms, we decided to do a carryover contamination experiment in an open platform,” Masciotra said.

‘Both algorithms performed similarly in early stages of infection before ART initiation.’
—Silvina Masciotra, MS

In comparing the two algorithms, the CDC used 46 U.S. seroconverters (subtype B), with 255 longitudinal samples before antiretroviral therapy initiation and 73 samples after ART initiation. “All of those were BioPlex seroreactive, and they were all positive after the first HIV-1 RNA,” she said. The comparison also included 105 Cameroonian ART-naïve specimens with established HIV-1 infection; three had HIV-1 group O infection and 102 had HIV-1 group M, non-B subtype infection. The CDC conducted the evaluations as part of collaborations with Bio-Rad and Hologic, both of which provided kits.

The BioPlex assay had a specificity of 99.7 percent. Aptima Quant had a specificity of 99.8 percent and no carryover contamination was observed.

The three-step algorithm detected HIV-1 infection in 96.2 percent of 79 samples from seroconverters with early infection before ART initiation. There was good agreement between the BioPlex and Geenius assays regarding HIV-1 detection, and HIV-2 reactivity was not observed with either assay, Masciotra said. Results were confirmed with the Aptima qualitative assay.

The two-test alternative algorithm detected 98.7 percent with a viral load range of less than 1.47 to more than 7 log (cop/mL). “We had two samples that were detected but non-quantified at this early stage of infection, and two of the Aptima qualitative assay nonreactives were either less than 1.47 or 4.89 log (cop/mL).” When Masciotra performed the McNemar’s comparison analysis (p=0.4795), there were no significant differences. “Both algorithms performed similarly in early stages of infection before ART initiation,” she said.

In later stages of infection, with the three-test algorithm, 176 samples had HIV-1 positive results on the seroconversion panels with the Geenius assay. All Cameroonian samples with established HIV-1 infection were Geenius HIV-1 positive. “One sample was also HIV-2 untypable where it showed HIV-2 reactivity,” she said. “We did not see HIV-2 reactivity on the BioPlex assay.” Further testing found no evidence of HIV-2 infection.

Comparison of the two algorithms on samples with established HIV-1 infection before ART initiation revealed a detection rate of 99.3 percent with the two-step algorithm, Masciotra said. The viral load range was from less than 1.47 to greater than 7 log (cop/mL). Two samples were nonreactive with the Aptima Quant and the Aptima qualitative assays, and seven samples were detected and not-quantified later in infection.

“When we looked at the McNemar’s comparison analysis, there was similar performance between the two algorithms in established infection before ART initiation,” Masciotra said.

Turning to detection of HIV-1 infection immediately after ART initiation, the three-test algorithm detected HIV-1 in 100 percent of the 73 samples. Nine samples from seroconverters showed seroreversion. “People go through viral suppression over time. You will see seroreversion to even an HIV antibody negative. But those were Aptima qualitative assay positive,” Masciotra said.

The two-test algorithm showed a lower detection rate of 87.7 percent. The viral load range extended from target not detected to 6.9 log (cop/mL). Aptima Quant did not detect three of the nine samples that seroreverted. Six of the 64 HIV-1 Geenius positive samples were target not detected on Aptima Quant.

“In this case, although the number of samples is limited, the McNemar’s analysis showed significant differences in the two-test algorithm,” Masciotra said. “It did not perform as well as the three-test algorithm.”

The study’s limitations were that there were not enough samples to repeat. The Aptima qualitative and quantitative assays were not performed in parallel for a set of ART-naïve seroconversion panels. The Geenius HIV 1/2 differentiation assay was performed using software version 1.1, which was prior to the update to address HIV-2 indeterminate results, though “there were no HIV-2 indeterminate results in our sample set,” she said. And there was a small number of samples from ART-treated persons.

Aptima Quant worked well for diagnosis and quantification as a second step in the proposed algorithm in different stages of HIV-1 infection, Masciotra said, though it is not FDA approved with a dual claim. The assay’s performance decreased after the IgG response is elicited and with suppressed viremia due to ART. “Maybe they can use an antibody test when there is an undetectable viral load result,” she said.

Confirmation with a dual-claim RNA assay is a plus for patient care, Masciotra said. “However, additional factors, such as the implications of an off-label use and costs associated with the implementation of a second-step NAT algorithm, need to be explored.”

Amy Carpenter Aquino is CAP TODAY senior editor.