EVALUATION OF AN IgM IMMUNOBLOT KIT FOR DENGUE DIAGNOSIS

GORO KUNO, C. BRUCE CROPP, JOLENE WONG-LEE, AND DUANE J. GUBLER
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado; Genelabs Technologies, Inc., Redwood City, California

Abstract. A commercial IgM immunoblot kit was evaluated for dengue diagnosis with a panel of serum specimens collected from patients in a dengue endemic area. The kit is not recommended for use in its present form because of its undesirable rate of false-positive results. However, by substituting internal controls with the reference positive and negative controls that are more representative of those seen in endemic areas and by modifying the positive and negative scoring criteria, sensitivity and specificity of 80.3% and 94.5%, respectively, were obtained. These results are comparable with those obtained with the IgM ELISA on specimens, most of which were obtained from outpatient health care facilities. With further technical modifications, inclusion of a visual guide to ensure scoring standardization, and a more complete elaboration of the limitations of the test, wide application of the kit in diagnostic laboratories should be possible.

Dengue fever is the most important arboviral disease in the world in terms of morbidity and mortality. Incidence of the severe forms of dengue, dengue hemorrhagic fever and dengue shock syndrome (DSS), which emerged in Southeast Asia in the 1950s and 1960s, has increased dramatically in the Americas. Since the first major DHF outbreak in Cuba in 1981, 23 countries have reported confirmed cases of DHF, and outbreaks have occurred in Venezuela, Brazil, Colombia, and French Guiana.

Although many diagnostic techniques are available, the major problems of dengue diagnosis in many countries in the tropics are high cost and lack of reagents and equipment. Since reagents for most arboviral diseases, including dengue, have not been commercially available until recently, diagnostic services have been limited to a very small number of well-established and equipped laboratories. This is one of the major reasons for poor surveillance, under-reporting of cases, and the long time intervals between the onset of epidemic transmission and the mobilization of mosquito control. For physicians administering medical care, a major problem is the long period between submission of specimens and receipt of results. The recent proliferation of commercial dengue diagnostic kits is highly encouraging in this regard. In particular, the kits that are simple to use are important for field applications in smaller clinics, hospitals, and other health care facilities. Recently, an immunoblot kit that satisfies some of those requirements was evaluated and judged to have excellent qualities in a multicenter study.

We have evaluated the same kit but obtained different results. Here, we report those data, discuss the problems of kit evaluation for dengue IgM detection, and present thoughts on how diagnostic kits should be evaluated.

MATERIALS AND METHODS

Clinical specimens. A panel of 233 serum specimens (96 paired specimens and 137 single specimens) were collected from dengue patients in Puerto Rico. Confirmation of infection was by virus isolation, by the standard hemagglutination inhibition (HI) test on paired serum specimens, and by detecting dengue-specific IgM antibody with the IgM antigen capture ELISA (MAC-ELISA). Ten cerebrospinal fluid (CSF) specimens from serologically confirmed Japanese encephalitis patients from the Philippines and nine serum spec-
er A was involved in arbovirus investigation and had experienced two preliminary blind test trials prior to this study; reader B was not involved in arbovirus investigation and had not participated in the preliminary trials. The readers were instructed first to read the blots by following the kit instructions, using intensities of reactive and nonreactive controls provided with the kit as guides; next, the readers were instructed to read the blots according to the modified criteria described below. All reagents were kept refrigerated before use, and processed membranes were kept in the dark (wrapped in aluminum foil) and refrigerated thereafter.

**Kit diagnostic criteria.** When a spot on the membrane showed a color intensity darker than that on nonreactive control and the reactive control showed a clear circular or ring image, the specimen was considered positive for IgM antibody; when the color intensity was the same or less than that of the nonreactive control, it was interpreted to be negative for IgM. It should be noted that the nonreactive control provided in the kit produced no visible image.

**In-house modified diagnostic criteria.** During two preliminary trials, we recognized the need to modify the criteria for more accurate interpretation of the blot. For these criteria, in-house positive and negative controls were used. On an arbitrary scale of − to +4, the positive and negative controls produced intensities of 4+ and 1+, respectively. A total of seven negative specimens, which produced ambiguous blot images (± or 1+), were included in each test. Readers were instructed to read those or similar specimens as negative and to score any specimens with intensities equal to or greater than 2+ as positive.

**IgM-capture ELISA and HI test.** The MAC-ELISA was used as reference and performed, as described earlier. The optical density (OD) values throughout the text represent the mean OD values of two wells per specimen after the mean OD value of the negative control specimen in each plate was subtracted. The standard HI test adapted for microtiter plate was performed, according to the method of Clarke and Casals. Dengue infections were classified as primary or secondary on the basis of the criteria established by the World Health Organization (WHO).

**Sensitivity and specificity.** A subset of 119 serum specimens obtained either in the acute or convalescent phases and previously tested for anti-dengue IgM by MAC-ELISA were classified as strong-positive (seven specimens), low-to-moderately positive (58 specimens), and negative (54 specimens) based on OD at 405 nm (OD₄₀₅ ≥ 0.60, OD₂₀₅ = 0.20–0.35, and OD₄₀₅ < 0.20, respectively). These specimens were selected solely on the basis of the OD reading in the MAC-ELISA and represented both acute phase (≤ 5 days after onset) and convalescent phase (≥ 8 days after onset). The IgM-negative specimens were further tested for flavivirus IgG by the HI test and found to be negative. Previously, a correlation was found to exist, within limits, between OD values and IgM titer for flaviviruses, except when OD values are very high; this has been used to extrapolate the IgM titer in units by adopting a fixed, cut-off OD ratio obtained with a reference positive specimen as 40 units. Thus, those two groups of IgM-positive specimens we tested roughly represented serum specimens of high and low to moderate IgM titers, respectively. This subset of specimens was used for comparative sensitivity/specificity tests between the two readers. In composing the panel of IgM-positive specimens for sensitivity and specificity testing, we arbitrarily selected low-to-moderately positive specimens to represent 89% of the positive samples because data on reliability of the IgM immunodot kit must be based on acute-phase samples, most of which are known to have low IgM titers in dengue infections. Furthermore, testing a large number of strongly positive specimens was unnecessary because in two preliminary trials all 12 strongly positive specimens produced intensely stained circles and the sensitivity was 100% in both trials.

**Statistical methods.** Data were analyzed statistically by the chi-square and McNemar tests.

### RESULTS

**Sensitivity and specificity.** According to reader A, who had previous reading experience during the preliminary trials, the sensitivity results of the immunoblot test using the kit and the modified criteria on a panel of 119 serum specimens were 96.9% (63 of 65) and 87.7% (57 of 65), respectively; the corresponding specificities were 33.3% (18 of 54) and 90.7% (49 of 54), respectively (Table 1). Reader B, who had no experience in reading immunoblot results, read 63 specimens (39 positive and 24 negative) of this panel of specimens composed so that low-to-moderately positive specimens represented nearly 90% of the positive specimens. According to reader B, the sensitivity results were 100% (39 of 39) and 66.7% (26 of 39) by the kit and the modified criteria, respectively; the corresponding specificity results...
were 41.7% (10 of 24) and 95.8% (23 of 24), respectively (Table 1).

When the kit criteria for a positive blot were used, the differences between the two readers were insignificant for both sensitivity and specificity; however, when modified criteria were used, the difference for specificity was significant ($P < 0.01$, by chi-square test). For each reader, differences between the two criteria were significant for both sensitivity and specificity (Table 1).

The difference in sensitivity and specificity between the two readers was further analyzed by the McNemar test on a subset of 39 IgM-positive and 24 IgM-negative specimens that were read by both readers (Table 1). Significant differences were observed between the two readers for sensitivity, when the modified reading criteria were used ($P < 0.05$, and between the two criteria for specificity by both readers ($P < 0.05$ and $P < 0.01$).

**Primary infection.** Among 20 pairs of specimens from primary infections, nine acute and 19 convalescent specimens had detectable IgM antibody by MAC-ELISA. The only IgM-negative convalescent specimen had HI titers (20–80) with the highest titer corresponding to DEN-1 virus. To compare the diagnostic accuracy of immunoblot with that of HI, if at least one of the paired specimens had detectable IgM, simply for facilitating comparison, we interpreted it as a correct diagnosis. When this interpretation was used, 18 pairs were correctly diagnosed by both readers with the kit criteria; and 19 and 17 pairs were diagnosed correctly by readers A and B, respectively, using the modified criteria (Table 2). The differences between readers or between reading criteria were not statistically significant ($P > 0.5$, by chi-square test).

**Secondary infection.** Of 20 serum pairs from patients with secondary infection, nine acute and 19 convalescent specimens were positive by the MAC-ELISA (Table 2). The only IgM-negative convalescent specimen had an HI titer of 2,560 for all four serotypes. Both readers correctly diagnosed 19 cases using the kit criteria; by the modified criteria, 17 and 16 cases were diagnosed correctly by readers A and B, respectively. There was no statistically significant difference between the two readers ($P > 0.5$, by chi-square test).

**Paired specimens with a stable HI titer ($\leq 640$).** Paired specimens from eight patients collected more than seven days apart were classified as not dengue on the basis of HI results, as defined by the WHO criteria. Five of the eight pairs had detectable IgM by the MAC-ELISA and four pairs had IgM by the immunoblot test. Overall, concordant results were obtained in six and seven cases, respectively, by the immunoblot using kit and modified criteria.

**Overall concordance of the immunoblot test.** Variation in generating concordant results (using the MAC-ELISA as a reference) between trials was analyzed for each reader. One hundred eighty specimens (which included all of the paired samples described above, plus 84 single specimens) were nearly equally divided for each category of specimens and used in two trials (Table 3). There were no statistically significant differences in concordance between the two readers in both trials ($P > 0.10$, by chi-square test). The membranes were stored at 4°C for one month, and read again by reader A; the same score was obtained.

**Flavivirus cross-reactivity.** Flavivirus cross-reactivity was extensive by the immunoblot test since eight of 10 CSF specimens of Japanese encephalitis and seven of nine yellow fever specimens were scored as positive, using both criteria by two readers.

**Variation among lots.** No obvious variation in quality was observed among five lots of the kit evaluated in this study.

**DISCUSSION**

Progress in developing new diagnostic techniques for viral diseases has been dramatic in the past two decades. Although technological improvements have been achieved for applications to field conditions, few techniques have met the four basic requirements needed for arboviral disease diagnoses in tropical countries: simplicity, economy, rapidity, and accuracy.

The immunoblot test is one of several variations of diagnostic techniques based on the visualization of a specific reaction with target molecules, such as antigen, antibody, and nucleic acid, on synthetic membranes or filter paper. Many of serologic techniques were designed to detect specific IgM in the acute phase of illness for early diagnosis. Thus, development of a similar test based on detection of specific dengue IgM was significant and deserves an objective evaluation of its qualities.

In evaluating diagnostic kits, it is most ideal to conduct a
multi-institutional study. Customarily, a panel of samples, including control samples, are aliquoted, coded, and distributed to all collaborating laboratories, and the results are interpreted blindly.\textsuperscript{14-17} Such an arrangement has been practiced for proficiency test of dengue serologic diagnosis in a collaborative program between the Centers for Disease Control and Prevention (CDC) and the Pan American Health Organization for the Caribbean and Latin American countries. However, the recent report of a multicenter evaluation of the same dengue IgM immunodot kit used in this study is unusual in that each participating laboratory evaluated the kit by using a panel of uncoded (except in one of six participating laboratories) samples selected independently in each laboratory.\textsuperscript{5} In an evaluation of another dengue diagnostic kit, a dipstick IgM ELISA kit, similarly no blind test was performed.\textsuperscript{18} With such an arrangement it is more difficult to draw a meaningful conclusion.

For evaluating the qualities of diagnostic tests that depend on subjective visual reading, unlike most other tests, selection and composition of a panel of specimens used affect the outcome. This is because, as observed in our study, visual perception for discriminating positive from negative specimens itself varies between readers and, most importantly, is influenced by the proportion of positive specimens and their titer. This problem was also evident in classifying the reactions of an immunochromatographic, diagnostic kit for dengue into four categories (no reactivity, faint, distinct, and strong) “depending on the intensity of the positive reaction.”\textsuperscript{19} Thus, when the positive specimens in a panel consist mostly of high-titered specimens, they yield intense color reactions in the blot test, resulting in very sharp contrast between those specimens and the negative control specimen; those positive specimens are almost always interpreted correctly, resulting in high sensitivity. On the other hand, when the majority of positive specimens in the panel are low-titered specimens, intensities of the color reaction are often weak, making the visual discrimination ambiguous even though positive and negative control images are used. In the latter case, sensitivity and/or specificity drop, as observed previously\textsuperscript{4} and in our study. Similarly, in an evaluation of the aforementioned immunochromatographic kit, nearly 50% of the reactions visually classified as faint or distinct were false positive.\textsuperscript{19}

Consequently, if a panel of positive specimens consists of a disproportionately large number of serum samples obtained in the late acute (\textgtrless\textequal{} 6 days after onset) or convalescent phase from primary dengue infections and from hospitalized patients, fewer false-positive or negative results are expected, since positive specimens in those groups of samples tend to have higher IgM titers. The overwhelming majority of dengue samples are from outpatients from whom acute phase specimens are usually all that is available. These samples generally have low IgM titers\textsuperscript{20,21} and produce ambiguous color of low intensities. For evaluation of test kits, therefore, it is important to use positive specimens with known IgM titer\textsuperscript{18,19} rather than just selecting any IgM-positive samples.

Our samples in the sensitivity and specificity study (Table 1) were arbitrarily drawn from both the acute and convalescent phases of illness, and do not represent a typical collection of blood samples obtained from hospitalized patients. We expect the IgM titers in most early acute-phase specimens to be low. Thus, inclusion of a large number of weakly or moderately IgM-positive specimens in the serum panel as we have done provides a more realistic measure of sensitivity of this immunoblot kit. For that reason, overall sensitivity of 80.3% scored by reader A, according to the modified criteria, compares favorably with 78% for a MAC-ELISA procedure\textsuperscript{10} or 71% on admission specimens alone using the immunochromatographic kit that detects both IgM and IgG.\textsuperscript{19}

In our study, specific IgM was not detectable in paired specimens of a small number of both primary and secondary infections, which confirms the observations in previous reports,\textsuperscript{22,23} and illustrates a limitation of IgM antibody tests for a small number of patients. Furthermore, detection of specific IgM in some cases, classified as “not current dengue,”\textsuperscript{18} raises a possibility that at least some of those cases demonstrating specific IgM in both acute- and convalescent-phase specimens collected nearly 2–3 months apart were current. On the other hand, some IgM ELISA-negative samples were immunoblot positive.\textsuperscript{5,13} The results could be interpreted as either evidence of superiority of the immunoblot technique or as false positive, if no other data are available for evaluation.

Our data suggest that the IgM blot kit needs improvement in the following ways before it is widely used. First, the nonreactive control (negative) needs to be replaced with a more representative negative control, which is prepared by pooling blood samples of febrile patients who were shown not to have dengue infection by reliable and commonly used laboratory tests. Second, the background in negative specimens may be further reduced by modifying reaction parameters, such as the optimal dilution of a serum specimen, the antigen preparation procedure, kind and concentration of antigen, and blocking. To improve the accuracy of the diagnosis, it is highly recommended that when a single acute-phase specimen gives an equivocal or negative result, that specimen should be considered negative until another properly timed specimen is collected and tested.

While the technical modifications noted above may improve the accuracy of the test, the most critical element for membrane-based diagnostic kits, such as immunoblot, Western blot, dipstick, and hybridization of DNA amplified by polymerase chain reaction,\textsuperscript{5,13,18,19,24-26} is subjectivity caused by variation in visual perception among diagnosticians. This is a critically important concern for the dengue kits in general because, unlike other diagnostic tests that can be performed only in well-established laboratories by experienced diagnosticians, the kits were designed for wider applications in field situations, such as small clinics, hospitals, and other health care facilities where test results may be interpreted by personnel without experience in dengue diagnosis. Although all dengue kits evaluated thus far were earlier recommended for use in peripheral institutions,\textsuperscript{5,18,19} we recommend against such use unless the diagnosticians responsible for interpreting the results are experienced in dengue diagnosis.

As revealed in our study, the variation in scoring among readers was evident on a subset of samples in which nearly 90% of positive samples were either weakly or moderately IgM positive. We observed that an individual with previous experience in interpreting blot results scored better than the other without any prior experience. Thus, the possibility remains that with a good panel of visual image controls and
training, reading can be further improved. Of interest is that a reader without prior experience (not included in our analysis) read the results of two trials using the modified criteria, improving sensitivity from 66.7% in the first trial to 77.8% in the second.

The high cross-reactivity observed in our test with CSF specimens from Japanese encephalitis patients and with serum samples from yellow fever patients is of concern. Cross-reactivity with serum samples from Japanese encephalitis patients was not reported with this immunoblot kit, but 50% of confirmed Japanese encephalitis specimens cross-reacted in another dengue kit. It was reported earlier that some IgM-positive CSF specimens of Japanese encephalitis infections were weakly positive by a dengue MAC-ELISA. Because the cases of central nervous system syndrome associated with dengue infection have been more frequently reported lately in tropical Asia where both dengue and JE are endemic, it is important that cross-reactivity information be clearly described. Also of concern is that the kit was reported to cross-react with malaria-positive specimens. Cross-reactivity with yellow fever IgM observed in this study was also reported in the evaluation of a dipstick IgM ELISA kit.

In summary, it is recommended that the immunoblot kit carry more complete information on its utility and limitations. Thus, the kit should not be used alone for diagnosis. In particular, it should not be used for managing hospitalized patients; it should be used only as an adjunct to clinical diagnosis. In addition, positive results by the kit cannot be regarded as an unequivocal laboratory diagnosis, but rather, should be interpreted only as evidence of recent flavivirus infection because of cross-reactivity and the persistence of IgM antibody. Furthermore, to better inform the users of the kit in peripheral health care facilities, it is recommended that two sets of sensitivity/specificity data on acute-phase specimens be provided with the kit: one set obtained using a related technique (such as the MAC-ELISA) as a reference and the other set showing the efficacy of that reference compared with a world standard (such as the HI test based on paired specimens). Despite the drawbacks above, if correctly used under circumstances in which the involvement of flaviviruses other than dengue are considered minimal or non-existent, this kit could be very useful in field settings. It is technically simple, does not depend on expensive equipment, requires very small sample volumes, and has reasonable sensitivity and specificity. Moreover, results can be obtained relatively rapidly.

Acknowledgments: Our appreciation is extended to Ivette Gomez (San Juan Laboratories, Division of Vector-Borne Infectious Diseases, CDC) for performing the HI test, and Raymond E. Bailey (formerly of CDC) for statistical assistance.

Disclaimer: Use of trade names is for identification only and does not imply endorsement by the Public Health Service or by the U.S. Department of Health and Human Services.

Authors’ addresses: Goro Kuno, C. Bruce Cropp, and Duane J. Gubler, Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, PO Box 2087, Fort Collins, CO 80525-2087. Jolene Wong-Lee, Genelabs Technologies, Inc., Redwood City, CA 94063.

REFERENCES


