

Evaluation of Two Commercial Systems for Automated Processing, Reading, and Interpretation of Lyme Borreliosis Western Blots[∇]

M. J. Binnicker,^{1*} D. J. Jespersen,¹ J. A. Harring,¹ L. O. Rollins,¹ S. C. Bryant,² and E. M. Beito¹

Division of Clinical Microbiology and Department of Laboratory Medicine and Pathology¹ and Division of Biostatistics,² Mayo Clinic and Mayo Clinic College of Medicine, Rochester, Minnesota 55905

Received 1 February 2008/Returned for modification 27 March 2008/Accepted 29 April 2008

The diagnosis of Lyme borreliosis (LB) is commonly made by serologic testing with Western blot (WB) analysis serving as an important supplemental assay. Although specific, the interpretation of WBs for diagnosis of LB (i.e., Lyme WBs) is subjective, with considerable variability in results. In addition, the processing, reading, and interpretation of Lyme WBs are laborious and time-consuming procedures. With the need for rapid processing and more objective interpretation of Lyme WBs, we evaluated the performances of two automated interpretive systems, TrinBlot/BLOTrix (Trinity Biotech, Carlsbad, CA) and BeeBlot/ViraScan (Viramed Biotech AG, Munich, Germany), using 518 serum specimens submitted to our laboratory for Lyme WB analysis. The results of routine testing with visual interpretation were compared to those obtained by BLOTrix analysis of MarBlot immunoglobulin M (IgM) and IgG and by ViraScan analysis of ViraBlot and ViraStripe IgM and IgG assays. BLOTrix analysis demonstrated an agreement of 84.7% for IgM and 87.3% for IgG compared to visual reading and interpretation. ViraScan analysis of the ViraBlot assays demonstrated agreements of 85.7% for IgM and 94.2% for IgG, while ViraScan analysis of the ViraStripe IgM and IgG assays showed agreements of 87.1 and 93.1%, respectively. Testing by the automated systems yielded an average time savings of 64 min/run compared to processing, reading, and interpretation by our current procedure. Our findings demonstrated that automated processing and interpretive systems yield results comparable to those of visual interpretation, while reducing the subjectivity and time required for Lyme WB analysis.

Lyme disease is a multisystem, tick-borne disease caused by the spirochete *Borrelia burgdorferi*. In 2006, the Centers for Disease Control and Prevention (CDC) reported 19,931 cases of Lyme disease in the United States (7), confirming that the disease continues to represent a significant public health threat. The clinical manifestations of early localized disease range from nonspecific sequelae, including malaise, myalgia, and lymphadenopathy, to more characteristic findings, such as erythema migrans (EM). In the absence of appropriate therapy, disease progression may lead to significant complications, including rheumatologic, neurologic, or cardiac manifestations (15, 16).

The diagnosis of Lyme borreliosis (LB) can be made clinically when patients from regions where the disease is endemic present with EM (5, 8, 18). However, in patients without EM but with objective clinical findings suggestive of disseminated LB, serologic testing is an important diagnostic approach. Appropriate serologic testing should follow the two-tier algorithm recommended by the CDC (6), consisting of initial testing with a sensitive screening assay (e.g., enzyme immunoassay) with positive or equivocal specimens to be tested by Western blot (WB) analysis. Current CDC criteria for WB interpretation recommend that ≥ 2 bands on the immunoglobulin M (IgM) WB or ≥ 5 bands on the IgG WB be present for the immunoblot to be considered positive (6, 9). Although WB is considered to be highly specific, current testing protocols in most clinical laboratories rely on visual reading and interpretation of

WB strips. These procedures require the laboratory technologist to visually compare band intensities on the patient strip to those of a weakly reactive control. This approach is labor-intensive, time-consuming, and subjective, allowing for potential intra- and interlaboratory variation in WB reading and interpretation. Previous studies analyzing the performance of LB serologic tests among testing laboratories have demonstrated significant variation in results, even for more objective methods, such as enzyme immunoassay (2, 3, 10). Therefore, given the inherent subjectivity in reading and interpreting WBs for diagnosis of LB (i.e., Lyme WBs), one would expect to observe significant variation in WB results, with potentially adverse effects on the laboratory diagnosis of Lyme disease and subsequent patient management decisions. Due to the need for more objective and consistent interpretation of Lyme WBs, we undertook a study to evaluate and compare two systems (TrinBlot/BLOTrix [Trinity BioTech, Carlsbad, CA] and BeeBlot/ViraScan [Viramed Biotech AG, Munich, Germany]) designed for automated processing, reading and interpretation of Lyme WBs. The goal of the present study was to determine whether automated systems yield comparable results to visual reading and interpretation while reducing the subjectivity and time required for Lyme WB analysis.

MATERIALS AND METHODS

Serum specimens. A total of 518 consecutive, unique serum specimens submitted to our reference laboratory for routine LB serologic testing between June and September 2007 were included in the study. The specimens were submitted without accompanying clinical information. The study protocol was reviewed and approved by the institutional review board of the Mayo Clinic.

In addition, two Lyme WB performance panels consisting of 55 clinically characterized and laboratory-characterized serum specimens were purchased from the CDC and Boston Biomedica, Inc. (BBI; West Bridgewater, MA).

* Corresponding author. Mailing address: Mayo Clinic, 200 First Street SW Hilton 860A, Rochester, MN 55905. Phone: (507) 538-1640. Fax: (507) 284-4272. E-mail: binnicker.matthew@mayo.edu.

[∇] Published ahead of print on 7 May 2008.

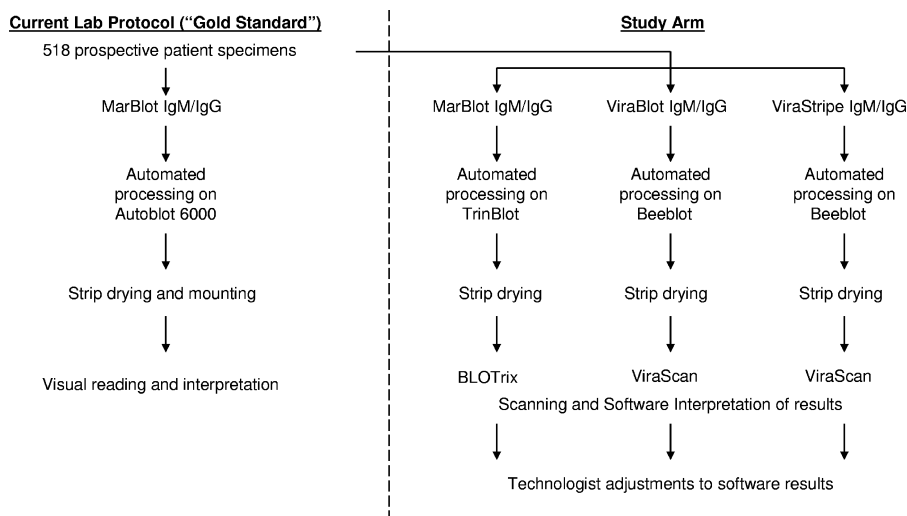


FIG. 1. Study design.

Study design. Each specimen was processed by our current procedure, performed according to the manufacturer's instructions for processing MarBlot IgG and IgM strips (MarDx Diagnostics, Carlsbad, CA) using the Autoblot 6000 instrument (MedTec, Inc., Hillsborough, NC). The strips were visually interpreted, and the results were recorded manually. Each specimen was also tested by MarBlot IgG and IgM strips using the automated TrinBlot processor (Bee Robotics, Caernarfon Gwynedd, United Kingdom) with subsequent scanning and analysis by the BLOTrix interpretive software. In addition, each specimen was tested by the ViraBlot and ViraStripe IgG and IgM strips (Viramed Biotech AG) on the automated BeeBlot processor (Bee Robotics) with subsequent analysis by the ViraScan interpretive software. A laboratory technologist blinded to the results of visual interpretation in the laboratory then reviewed the BLOTrix and ViraScan software interpretive results for each specimen (Fig. 1). The laboratory technologist, when reviewing the software interpretive results for each strip, (i) ensured that the software had analyzed only bands demonstrating uniform intensity across the entire width of the strip, (ii) checked that any background intensity (non-band intensity) had been accounted for, (iii) verified that the software had properly aligned test bands on the patient strip with bands on the serum band locator control strip, and (iv) ensured that the densitometric read was focused on the center of each test band.

WB assays. MarDx MarBlot IgM and IgG assays (Fig. 2) utilize antigens of *B. burgdorferi* strain B31 for Western blot analysis. The antigens are separated by electrophoresis through a sodium dodecyl sulfate-polyacrylamide gel. The resolved antigens are then transferred to a nitrocellulose membrane. Similarly, the ViraBlot and ViraStripe IgM and IgG assays use strain *B. burgdorferi* B31 as the source of antigen for Western blot analysis. ViraBlot IgM and IgG assays (Fig. 2) are manufactured by separation of antigens by gel electrophoresis, with subsequent transfer to nitrocellulose. In contrast, the ViraStripe IgM and IgG assays (Fig. 2) are generated by "printing" highly purified antigens at a defined location with standardized concentrations on a nitrocellulose membrane. The ViraStripe IgM and IgG assays are currently prototype Lyme WB strips, while the MarBlot and ViraBlot assays are both Food and Drug Administration cleared.

WB strip processing, reading, and interpretation. For each specimen tested by our current Lyme WB procedure, 80 μ l of the serum band locator, 20 μ l of the weakly reactive and negative controls, and 20 μ l of patient serum were added to the appropriate channels of the Autoblot 6000 instrument. After being incubated and washed, strips were air dried and mounted for visual reading and interpretation. Each test band was visually compared to the 41-kDa band on the weakly reactive control strip and was considered present if the intensity was equal to or greater than that of the weakly reactive control. The IgM assay was considered positive if two of the following three bands were present: 23, 39, and 41 kDa. The IgG assay was considered positive if five of the following ten bands were present: 18, 23, 28, 30, 39, 41, 45, 58, 66, and 93 kDa.

Each specimen was also tested by the ViraBlot and ViraStripe IgM and IgG assays, performed according to the manufacturer's instructions on the BeeBlot automated processor. In brief, strips were added to the incubation tray and allowed to presoak in 1.5 ml of wash buffer for 5 min. Next, 20 μ l of patient serum or 100 μ l of control was added to the appropriate channel of the incubation

tray. After being incubated and washed, strips were air dried in the incubation tray, scanned by using the ViraCam scanner (Viramed Biotech, AG), and analyzed by the ViraScan analysis software version 2.01. Based on densitometric analysis as described by Nishizuka et al. (12), 8-bit, grayscale images at a resolution of 220 dots per inch were stored in bitmap format for subsequent analysis by ViraScan. With the aid of predefined band-locator images, ViraScan locates and measures the immunospecific banding patterns on each patient strip. Each band is then analyzed for band location and maximum intensity. After a subtraction of background intensity (the average non-band intensity), the software assigns a numerical value to each band. The software then divides the numeric value of each test band by the numeric value assigned to a separate calibrator control band to determine a relative intensity. For the present study, a test band was considered present if the relative intensity met or exceeded the following manufacturer's recommended cutoff settings: ViraBlot IgG, 75%; ViraBlot IgM, 70%; ViraStripe IgG, 85%; and ViraStripe IgM, 60%.

Each specimen was also tested by the MarDx MarBlot IgM and IgG assays, performed according to the manufacturer's instructions using a TrinBlot automated processor. After automated processing, the strips were air dried in the incubation tray, scanned, and analyzed by the BLOTrix analysis software version 2.6. Similar to ViraScan, BLOTrix software utilizes densitometric analysis to compare the intensity of each test band to that of a separate calibrator control band and calculate a relative percent intensity. For the present study, a test band was considered present if the calculated relative intensity met or exceeded the following manufacturer's recommended cutoff settings: TrinBlot IgG, 90%; and TrinBlot IgM, 90%.

Statistics. Statistical analyses were performed by using statistical analysis software (SAS Institute Inc, Cary, NC). In addition to the percent agreement, kappa coefficients were determined as a secondary measure of agreement. Result agreements by kappa values are categorized as near perfect (0.81 to 1.0), substantial (0.61 to 0.80), moderate (0.41 to 0.60), fair (0.21 to 0.40), slight (0 to 0.20), or poor (<0) (11).

Workflow analysis. The average assay time for testing by our current procedure was calculated by timing three separate runs (40 specimens/run) from the addition of strips to the incubation tray through the manual recording of results by the technologist. The average assay time for testing by the automated systems was calculated by timing three separate runs (40 specimens/run) from the addition of strips to the incubation tray through the review of software results by the laboratory technologist.

RESULTS

Agreement between automated and visual interpretation. To assess agreement, the qualitative results (positive or negative based on CDC criteria) were compared after the testing of 518 consecutive serum specimens. Among the 518 specimens, 74 (14.3%) were positive for IgG, and 191 (36.9%) were positive

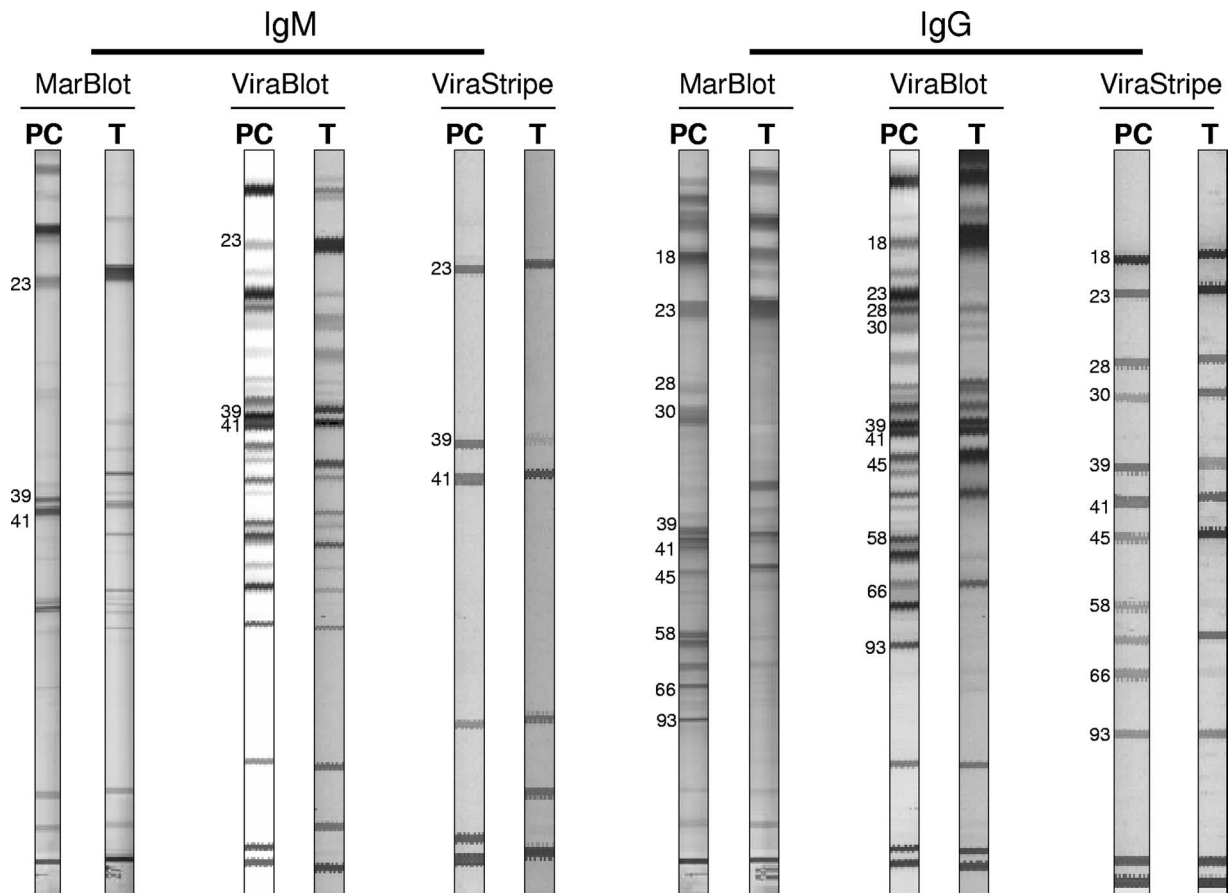


FIG. 2. Comparison of WB strips for a single patient specimen. The same patient specimen was tested by the MarBlot, ViraBlot, and ViraStripe assays. Strips were scanned by their respective systems, and the images were captured in tag image file format (TIFF). The migration positions of bands used in the CDC interpretation criteria are indicated by molecular mass (in kilodaltons). PC, positive control; T, test (patient) sample.

for IgM by our current procedure. BLOTriX analysis of the MarBlot assays demonstrated an agreement of 84.7% (439/518) for IgM and 87.3% (452/518) for IgG compared to results obtained by routine testing (Table 1). ViraScan analysis of the ViraBlot IgM and IgG assays showed 85.7% (444/518) and

94.2% (488/518) agreement, respectively, while analysis of the ViraStripe assays demonstrated 87.1% agreement (451/518) for IgM and 93.1% agreement (482/518) for IgG (Table 1). A technologist blinded to the results of routine testing then reviewed the BLOTriX and ViraScan interpretive results for each

TABLE 1. Agreement of Lyme WB strip results obtained by visual interpretation, automated software, and technologist-adjusted software interpretation among prospective serum specimens ($n = 518$)^a

Assay ^b	Software alone		Technologist adjusted		K ^d
	% Agreement (95% CI)	K ^c	% Agreement (95% CI)	K ^c	
MarBlot					
IgM	84.7 (80.3–88.6)	0.69	87.1 (82.6–90.9)	0.73	0.86
IgG	87.3 (84.0–90.0)	0.57	92.3 (89.0–95.0)	0.72	0.77
ViraBlot					
IgM	85.7 (81.3–89.6)	0.70	85.9 (81.5–89.8)	0.71	0.97
IgG	94.2 (90.9–96.9)	0.75	94.8 (90.5–96.5)	0.75	0.90
ViraStripe					
IgM	87.1 (82.6–90.9)	0.72	86.7 (82.3–90.6)	0.71	0.97
IgG	93.1 (89.7–95.7)	0.72	91.7 (88.4–94.4)	0.69	0.89

^a Percent agreement refers to percent agreement with visual interpretation of MarDx MarBlot strips.

^b MarBlot results were analyzed by using BLOTriX software application; ViraBlot and ViraStripe results were analyzed by using the ViraScan software application.

^c Kappa coefficient for comparison with visual interpretation.

^d Kappa coefficient, software-alone interpretation versus technologist-adjusted software interpretation.

TABLE 2. Agreement of Lyme WB strip results obtained by visual interpretation and technologist-adjusted software interpretation among the BBI and CDC serum performance panels

Panel	% Agreement (95% CI) ^a			
	MarBlot*	MarBlot†	ViraBlot	ViraStripe
BBI				
IgM	73.3 (46.0–91.1)	73.3 (46.0–91.1)	93.3 (62.0–100.0)	80.0 (51.3–96.4)
IgG	86.7 (57.0–100.0)	86.7 (57.0–100.0)	86.7 (57.0–100.0)	80.0 (51.7–96.1)
CDC				
IgM	NA	80.0 (62.6–92.1)	67.5 (51.2–80.7)	82.5 (64.9–94.4)
IgG	NA	90.0 (71.8–100.0)	92.5 (74.1–100.0)	90.0 (71.8–100.0)

^a The percent agreement with reference WB results provided by BBI or the CDC. *, tested by the current procedure with visual reading and interpretation; †, analyzed by using the BLOTrix software application. ViraBlot and ViraStripe results were analyzed by using the ViraScan software application. NA, not applicable (inadequate specimen volume to be tested by current lab procedure with visual interpretation).

specimen. The technologist-adjusted BLOTrix and ViraScan results showed substantial agreement ($0.61 < \kappa < 0.80$) with results obtained by routine testing with visual interpretation (Table 1).

In addition to the analysis of 518 consecutive serum specimens, two performance panels (BBI and CDC) consisting of 55 serum specimens were tested by the automated systems. BBI serum specimens were also tested by our current Lyme WB procedure. Agreement was assessed by comparing the results of testing to the reference WB results provided with the performance panels (Table 2). Interestingly, the technologist-adjusted ViraScan results of the ViraBlot and ViraStripe IgM assays showed closer agreement (93.3 and 80%, respectively) with the reference BBI WB results than did the results obtained by routine testing (73.3%) (Table 2).

Sensitivity and specificity of automated systems. After the testing of 518 prospective specimens, the results were analyzed by using the visual interpretation of the MarDx assays as the “gold standard.” Although the automated systems demonstrated comparable performances overall, BLOTrix analysis showed higher sensitivities for IgM and IgG (93.3 and 81.1%, respectively) than did ViraScan analysis of the ViraBlot (86.1 and 74.3%, respectively) or ViraStripe (78.9 and 77.0%, respectively) strips (Table 3). In contrast, ViraScan analysis of the ViraStripe IgM and IgG strips showed the highest speci-

ficity (92.0 and 95.7%, respectively) compared to visual interpretation of the MarDx MarBlot assays (Table 3).

Adjusted sensitivities and specificities were then calculated after visual review of the automated software results by a laboratory technologist (Table 3). A total of 37/518 (7.1%) MarBlot IgM and 40/518 (7.7%) MarBlot IgG results were adjusted after visual review of the BLOTrix analyses. The majority of the changes ($56/77 = 72.7\%$) made to the BLOTrix interpretations were from positive to negative. These adjustments yielded a marginal increase in specificity for the MarBlot assays (Table 3). In contrast, a total of 8/518 (1.5%) ViraBlot IgM, 12/518 (2.3%) ViraBlot IgG, 8/518 (1.5%) ViraStripe IgM, and 15/518 (2.9%) ViraStripe IgG results were adjusted after review of the ViraScan analyses. The majority of changes ($35/43 [81.4\%]$) made to the ViraScan interpretations were from negative to positive, resulting in the adjusted sensitivities and specificities outlined in Table 3.

The clinical sensitivity and specificity of the automated systems were further evaluated by using the 40-member CDC serum performance panel. Specimens were categorized by clinical diagnosis according to detailed histories included with the panel, and WB results were analyzed by comparison to the clinical findings. Reference CDC WB results showed a sensitivity of 51.4% (18/35) for IgM and 48.6% (17/35) for IgG in patients with confirmed or probable Lyme disease (Table 4).

TABLE 3. Sensitivity and specificity of automated software or technologist-adjusted software interpretation in prospective serum specimens ($n = 518$)^a

Assay ^b	Sensitivity (95% CI)		Specificity (95% CI)		No. of technologist adjustments (% of total tests)
	Software alone	Technologist adjusted	Software alone	Technologist adjusted	
MarBlot					
IgM	93.3 (88.9–96.0)	92.8 (88.3–95.7)	79.6 (75.8–82.8)	83.6 (79.2–87.3)	37 (7.1)
IgG	81.1 (70.7–88.4)	86.5 (76.9–92.5)	88.3 (85.0–91.0)	93.2 (90.5–95.2)	40 (7.7)
ViraBlot					
IgM	86.1 (82.3–88.5)	86.1 (82.3–88.5)	85.5 (81.2–88.9)	85.8 (82.4–88.4)	8 (1.5)
IgG	74.3 (66.3–82.9)	81.1 (70.7–88.4)	97.5 (95.6–98.6)	96.0 (93.7–97.4)	12 (2.3)
ViraStripe					
IgM	78.9 (72.6–84.0)	77.8 (73.7–80.9)	92.0 (88.5–94.5)	92.0 (89.1–93.9)	8 (1.5)
IgG	77.0 (66.3–85.1)	81.1 (70.7–88.4)	95.7 (93.4–97.2)	93.5 (90.8–95.4)	15 (2.9)

^a Visual interpretation of MarDx MarBlot strips served as the gold standard for sensitivity and specificity results.

^b MarBlot results were analyzed by using the BLOTrix software application; ViraBlot and ViraStripe results were analyzed by using the ViraScan software application.

TABLE 4. Clinical sensitivity and specificity of technologist-adjusted software interpretation among the CDC serum performance panel ($n = 40$)

Clinical diagnosis	No. of patients	No. (%) positive ^a by:							
		CDC reference WB*		MarBlot (BLOTrix)†		ViraBlot (ViraScan)†		ViraStripe (ViraScan)†	
		IgM	IgG	IgM	IgG	IgM	IgG	IgM	IgG
Confirmed Lyme disease ^b	28	16 (57.1)	10 (35.7)	12 (42.9)	6 (21.4)	27 (96.4)	8 (28.6)	10 (35.7)	11 (39.3)
Probable Lyme disease ^c	7	2 (28.6)	7 (100.0)	2 (28.6)	7 (100.0)	3 (42.9)	6 (85.7)	1 (14.3)	6 (85.7)
Subtotal	35	18	17	14	13	30	14	11	17
Healthy donor	5	0 (0)	0 (0)	0 (0)	0 (0)	1 (20.0)	0 (0)	0 (0)	0 (0)
Clinical sensitivity ^d (%)		51.4	48.6	40.0	37.1	85.7	40.0	31.4	48.6
Clinical specificity ^e (%)		100.0	100.0	100.0	100.0	80.0	100.0	100.0	100.0

^a *, MarDx MarBlot (visual interpretation at the CDC); †, results obtained after technologist-adjusted software interpretation.

^b Physician-documented EM and/or positive culture for *Borrelia burgdorferi*.

^c Positive LB serology and meeting CDC case definition.

^d The subtotal number compared to the total number of confirmed and probable Lyme disease cases ($n = 35$).

^e The number of negative healthy donor results compared to the total number of healthy donors ($n = 5$).

The technologist-adjusted BLOTrix results showed a sensitivity of 40.0% (14/35) for IgM and a sensitivity of 37.1% (13/35) for IgG. In contrast, ViraScan analysis of the ViraBlot IgM and IgG strips demonstrated sensitivities of 85.7% (30/35) and 40.0% (14/35), respectively, while the ViraStripe assays showed sensitivities of 31.4% (11/35) for IgM and 48.6% (17/35) for IgG. Each of the systems demonstrated 100% specificity for IgM and IgG with the exception of the ViraBlot IgM assay, which showed a specificity of 80% (4/5) (Table 4).

DISCUSSION

It is estimated that over 2.5 million LB serology tests are performed annually in the United States (1, 18). In 2006, our laboratory at the Mayo Clinic performed 75,478 LB serology tests, with 37,338 (49.5%) of these being done by Lyme WBs. These numbers indicate that LB serology continues to play an important role in the diagnosis of the disease. Furthermore, these data emphasize the need to improve the efficiency of Lyme WB processing, reading, and interpretation due to the significant time and effort required by laboratory personnel.

A significant limitation of current Lyme WB testing is the subjectivity involved in the visual reading and interpretation of test strips. Preliminary studies in our laboratory have demonstrated considerable variation in Lyme WB results when strips are visually read by different laboratory technologists (M. Binnicker, unpublished data). This variation in WB results may contribute to inaccurate diagnoses, resulting in various consequences to patients, as described in past studies (4, 13, 14, 17). Therefore, an important need of clinical laboratories is to enhance the objectivity and consistency of Lyme WB interpretation.

Our findings show substantial agreement between the results of automated and visual interpretation of Lyme WBs. Both systems we evaluated demonstrated comparable results, excellent reproducibility (data not shown), and similar features and total average assay times (Table 5). We should emphasize that both systems are designed to aid in band identification and result interpretation and yet require a laboratory technologist to review and verify results prior to reporting. In our experience, the ViraScan software application was more intuitive to operate and required fewer result modifications by the review-

ing laboratory technologist (Table 3). This difference may be due, in part, to the specific manufacturer's recommended cutoff settings used in our evaluation. Clinical laboratories should perform their own thorough evaluation prior to implementing an automated system, since the appropriate cutoff settings may differ between regions where Lyme disease is and is not endemic.

The present study has several additional limitations. First, the number of clinically characterized specimens tested was limited and, therefore, no firm conclusions can be made regarding the accuracy of the automated interpretive systems. Future studies should test a large panel of clinically defined and laboratory-defined specimens in order to more accurately determine the clinical sensitivity and specificity. Second, the data presented here compare the qualitative interpretations (positive versus negative) using the current CDC criteria and do not focus on a direct comparison of individual bands. However, we observed very good correlation, overall, for the detection of specific bands between automated and visual interpretation. Interestingly, correlation seemed to be lowest for the 41- and 58-kDa bands (data not shown), and this may be due, in part, to their migration proximity to the 39- and 60-kDa

TABLE 5. Comparison of features between automated systems

System feature	TrinBlot/BLOTrix	BeeBlot/ViraScan
Software application	BLOTrix	ViraScan
Total assay time/run (h) ^a	2.62	2.65 ^b
Maximum run size ^c	50	50
Scanning resolution (dpi) ^d	104	220
IgG and IgM in single run	No	Yes
Electronic interface with LIS ^e	Yes	Yes
Serum and conjugate control bands on assay strip	Yes	Yes

^a This result represents the average time of three separate runs (adding strips to the incubation tray through technologist review and verification of software results; 40 specimens/run). The total assay time for the current protocol using MarBlot strips (adding strips to the incubation tray through visual reading and interpretation; 40 specimens/run) was 3.7 h.

^b The total assay time for either ViraBlot or ViraStripe.

^c The maximum run size for the current protocol (MarBlot strips on AutoBlot 6000) = 60 strips.

^d dpi, dots per inch.

^e LIS, Laboratory Information System.

antigens, respectively. These antigens are often difficult to distinguish by visual interpretation and may also require a more detailed verification following the automated software interpretation. A third limitation of the present study is that the same laboratory technologist reviewed and verified the software results for each specimen. Although this approach was used for consistency, the decision to manually adjust software results may vary between technologists. In order to minimize subjectivity and result variability, it will be essential for testing laboratories to establish specific criteria to guide and regulate the modification of software results. A point of interest for future studies will be to determine whether automated systems decrease inter- and intralaboratory result variability in comparison to visual reading and interpretation.

In summary, automated systems showed results comparable to those obtained by routine testing, while demonstrating several advantages. First, the average turnaround time was reduced by 64 min/run, translating into a time savings of approximately 1,000 h/year for clinical laboratories testing 37,000 to 40,000 specimens by WBs. Second, automated systems yield an approximate savings of 0.3 full-time equivalent in comparison to routine testing with visual interpretation. Additional benefits include the ability to electronically store data, share results with clinicians, and interface results with the Laboratory Information System. Finally, automated systems allow for a more objective interpretation of test strips, which may prove to significantly enhance the consistency of Lyme WB results.

ACKNOWLEDGMENTS

We thank the laboratory technologists and assistants in the Infectious Diseases Serology laboratory at the Mayo Clinic, who provided excellent laboratory and technical support during this study. We also thank Joseph Yao for critical review of the manuscript. The reagents and kits used in this study were provided by Trinity Biotech and Viramed Biotech AG.

REFERENCES

1. **Aguero-Rosenfeld, M. E., G. Wang, I. Schwartz, and G. P. Wormser.** 2005. Diagnosis of Lyme borreliosis. *Clin. Microbiol. Rev.* **18**:484–509.
2. **Bakken, L. L., K. L. Case, S. M. Callister, N. J. Bourdeau, and R. F. Schell.** 1992. Performance of 45 laboratories participating in a proficiency testing program for Lyme disease serology. *JAMA* **268**:891–895.
3. **Bakken, L. L., S. M. Callister, P. J. Wand, and R. F. Schell.** 1997. Interlaboratory comparison of test results for detection of Lyme disease by 516 participants in the Wisconsin State Laboratory of Hygiene/College of American Pathologists proficiency testing program. *J. Clin. Microbiol.* **35**:537–543.
4. **Brown, S. L., S. L. Handen, and J. J. Langone.** 1999. Role of serology in the diagnosis of Lyme disease. *JAMA* **281**:62–66.
5. **Bunikis, J., and A. G. Barbour.** 2002. Laboratory testing for suspected Lyme disease. *Med. Clin. N. Am.* **86**:311–340.
6. **CDC.** 1995. Recommendations for test performance and interpretation from the Second National Conference on the Serologic Diagnosis of Lyme Disease. *MMWR Morb. Mortal. Wkly. Rep.* **44**:590–591.
7. **CDC.** 2006. Reported cases of Lyme disease by year, United States, 1991–2006. Centers for Disease Control and Prevention, Atlanta, GA.
8. **Depietropaolo, D. L., J. H. Powers, J. M. Gill, and A. J. Foy.** 2005. Diagnosis of Lyme disease. *Am. Fam. Physician* **72**:297–304.
9. **Dressler, F., J. A. Whalen, B. N. Reinhardt, and A. C. Steere.** 1993. Western blotting in the serodiagnosis of Lyme disease. *J. Infect. Dis.* **167**:392–400.
10. **Hedberg, C. W., M. T. Osterholm, K. L. MacDonald, and K. E. White.** 1987. An interlaboratory study of antibody to *Borrelia burgdorferi*. *J. Infect. Dis.* **155**:1325–1327.
11. **Landis, J. R., and G. G. Koch.** 1977. The measurement of observer agreement for categorical data. *Biometrics* **33**:159–174.
12. **Nishizuka, S., N. R. Washburn, and P. J. Munson.** 2006. Evaluation method of ordinary flatbed scanners for quantitative density analysis. *BioTechniques* **40**:442–448.
13. **Reid, M. C., R. T. Schoen, J. Evans, J. C. Rosenberg, and R. I. Horwitz.** 1998. The consequences of overdiagnosis and overtreatment of Lyme disease: an observational study. *Ann. Intern. Med.* **128**:354–362.
14. **Sigal, L. H.** 1998. Pitfalls in the diagnosis and management of Lyme disease. *Arthritis Rheum.* **41**:195–204.
15. **Steere, A. C.** 1989. Lyme disease. *N. Engl. J. Med.* **321**:586–596.
16. **Steere, A. C., S. E. Malawista, D. R. Snyderman, R. E. Shope, W. A. Andiman, M. R. Ross, and F. M. Steele.** 1977. Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. *Arthritis Rheum.* **20**:7–17.
17. **Steere, A. C., E. Taylor, G. L. McHugh, and E. L. Logigian.** 1993. The overdiagnosis of Lyme disease. *JAMA* **269**:1812–1816.
18. **Tugwell, P., D. T. Dennis, A. Weinstein, G. Wells, B. Shea, G. Nichol, R. Hayward, R. Lightfoot, P. Baker, and A. C. Steere.** 1997. Laboratory evaluation in the diagnosis of Lyme disease. *Ann. Intern. Med.* **127**:1109–1123.