

Devices for Rapid Diagnosis of Malaria: Evaluation of Prototype Assays That Detect *Plasmodium falciparum* Histidine-Rich Protein 2 and a *Plasmodium vivax*-Specific Antigen

J. Russ Forney,^{1*} Chansuda Wongsrichanalai,² Alan J. Magill,³ Leslie G. Craig,¹ Jeeraphat Sirichaisinthop,⁴ Christian T. Bautista,³ R. Scott Miller,⁵ Christian F. Ockenhouse,⁶ Kent E. Kester,⁶ Naomi E. Aronson,⁵ Ellen M. Andersen,³ Higinio A. Quino-Ascurra,⁷ Carlos Vidal,⁷ Kimberly A. Moran,⁵ Clinton K. Murray,⁸ Caroline Cady DeWitt,⁹ D. Gray Heppner,⁶ Kevin C. Kain,¹⁰ W. Ripley Ballou,⁶ and Robert A. Gasser Jr.⁶

United States Military Academy, West Point, New York¹; Armed Forces Research Institute of Medical Sciences, Bangkok,² and Vector Borne Disease Control Office 1, Phrabuddhabat, Saraburi,⁴ Thailand; Naval Medical Research Center Detachment, Lima,³ and Hospital de Apoyo, Iquitos,⁷ Peru; Walter Reed Army Medical Center, Washington, D.C.⁵; Walter Reed Army Institute of Research, Silver Spring, Maryland⁶; The Toronto General Hospital and University of Toronto, Toronto, Ontario, Canada¹⁰; and Brooke Army Medical Center, San Antonio,⁸ and Wilford Hall Medical Center, Lackland Air Force Base,⁹ Texas

Received 6 January 2003/Returned for modification 13 February 2003/Accepted 17 March 2003

The *ParaSight F* test was developed as a pioneer industry effort in the large-scale, process-controlled production of a device for the rapid diagnosis of malaria. This device performed well in field settings but was limited to the detection of a single malaria species, *Plasmodium falciparum*. The *ParaSight F+V* assay advanced upon the *ParaSight F* test format by incorporating a monoclonal antibody directed against a proprietary *Plasmodium vivax*-specific antigen, in addition to the antibody directed against *P. falciparum* histidine-rich protein 2, which was used in the *ParaSight F* assay. The modified assay was developed to add the capability to detect *P. falciparum* and *P. vivax* in a single-test-strip format. The present study evaluated three distinct *ParaSight F+V* prototypes with samples from symptomatic patients in regions of Thailand and Peru where malaria is endemic. Over a 2-year enrollment period (1998 and 1999), a total of 4,894 patients consented to participation in the study. Compared with the results for duplicate microscopic examinations of Giemsa-stained blood smears as the reference diagnostic standard, each successive prototype showed substantial improvement in performance. The final *ParaSight F+V* prototype, evaluated in 1999, had an overall sensitivity for detection of asexual *P. falciparum* parasites of 98%. The sensitivity of the device was 100% for *P. falciparum* densities of >500 parasites/ μ l, with a sensitivity of 83% for parasite densities of \leq 500/ μ l. The specificity for the exclusion of *P. falciparum* was 93%. For *P. vivax*, the overall sensitivity was 87% for the final 1999 prototype. The sensitivities calculated for different levels of *P. vivax* parasitemia were 99% for parasite densities of >5,000/ μ l, 92% for parasite densities of 1,001 to 5,000/ μ l, 94% for parasite densities of 501 to 1,000/ μ l, and 55% for parasite densities of 1 to 500/ μ l. The specificity for the exclusion of *P. vivax* was 87%. The areas under the receiver operating characteristic curves for the diagnostic performance of the assay for the detection of *P. falciparum* and *P. vivax* were 0.8907 and 0.8522, respectively. These findings indicate that assays for rapid diagnosis have the potential to enhance diagnostic capabilities in those instances in which skilled microscopy is not readily available.

The limitations associated with diagnostic microscopy are well documented. These limitations have fostered the development of nonmicroscopic alternatives for the diagnosis of malaria (1, 3, 6, 11, 20, 21, 23, 26, 36). Devices for the rapid diagnosis of malaria (MRDDs [malaria rapid diagnostic devices]) emerged in the past decade to meet the need for a reliable diagnostic adjunct to microscopy in the clinical setting (40). Timely, accurate diagnosis has the potential to greatly facilitate the provision of appropriate therapeutic interventions, and as

the MRDD technology matures, these devices will have a pivotal role in malaria control initiatives.

The leading MRDD technology is the immunochromatographic strip (ICS) format, into which antigen-capture immunoassay methodologies are integrated. One of the earliest successful MRDDs, the *ParaSight F* test (Becton Dickinson Diagnostic Systems, Cockeysville, Md.) (32), used monoclonal-polyclonal antibodies directed against *Plasmodium falciparum* histidine-rich protein 2 (HRP-2) immobilized on a nitrocellulose strip. The *ParaSight F* test strip has been evaluated in a number of field studies and typically demonstrated clinically useful sensitivities and specificities (2, 4, 5, 7, 8, 10, 12, 15, 25; I. Traore, O. Koita, and O. Doumbo, *Am. J. Trop. Med. Hyg.*, abstr. 502, p. 272, 1997). The initial capabilities of MRDDs were limited to the detection of *P. falciparum*, i.e., a single-

* Corresponding author. Mailing address: ATTN: MADN-C&LS (COL Forney), United States Military Academy, Official Mail & Distribution Center, 646 Swift Rd., West Point, NY 10996-1905. Phone: (845) 938-5815. Fax: (845) 938-2235. E-mail: mj9173@usma.edu.

species format. Second-generation MRDD prototypes have retained the ability to distinguish this most clinically significant cause of malaria, while they have added the capability to identify other *Plasmodium* species (16, 17, 28, 33, 34, 37). Such refinement offers an enhanced potential for clinical utility and commercial success.

The growing number of published field studies of MRDDs for the detection of multiple species display wide degrees of diversity in their designs and methodologies. These differences significantly hinder interstudy comparisons. Protocols are designed and executed with different objectives. The characteristics of the enrolled populations vary, often substantially. Many field studies fail to document rigorous standards for reference microscopy, a particularly disconcerting lapse, given the critical role of microscopy in determining true diagnostic outcomes. Study populations are generally limited in size, and studies with small populations lack the statistical power to discriminate significant differences in device performance. The latter consideration limits studies in their ability to correlate performance characteristics with defined levels of parasitemia. Furthermore, it is frequently unclear which prototype of a device has been evaluated in a study and under what controls the manufacturing process was regulated. The continual evolution of nonmicroscopic assays for the diagnosis of malaria has led to the availability of multiple test prototypes, often indistinguishable on the basis of packaging, labeling, or production date. Inconsistencies in the reported performances of devices bearing common trade names may be attributable, at least in part, to variances in manufacturing and modifications inherent in iterative product development. These issues of protocol design and control, variance in production quality, and prototype identity have major implications for interpretations of device performance.

This report presents data for three distinct *ParaSight* F+V prototype assays that detect both *P. falciparum* HRP-2 and a *Plasmodium vivax*-specific antigen in a single ICS assay. These prototypes were products of the manufacturer's internal research and development initiatives. The assays differed in systematic modifications to assay materials, reagent formulations, and device assembly. All three prototypes were manufactured under process-validated controls in a manner consistent with U.S. Food and Drug Administration (FDA) guidelines for the production of in vitro diagnostic devices. The study protocol was specifically designed to evaluate device performance with samples from large populations of symptomatic patients enrolled from geographically distinct sites and to use a rigorously defined and controlled standard for the performance of diagnostic microscopy.

MATERIALS AND METHODS

Participants. Study participants were symptomatic patients who presented on their own initiative to existing outpatient malaria clinics. These malaria clinics, in Iquitos, Peru, and Maesot, Thailand, were operated by local public health authorities. During the study, the staff of these clinics retained full responsibility for patient care; all diagnostic and treatment decisions were independent of the study protocol. Patient enrollment was conducted in two separate increments, 28 May through 28 August 1998 and 17 May through 9 July 1999, during the peak malaria transmission seasons. The inclusion criteria were similar to those used in a previous study (10), as follows: (i) fever (oral temperature, $\geq 38^{\circ}\text{C}$), (ii) history of fever over the past 72 h, or (iii) headaches. In 1998, patients providing a history of antimalarial drug therapy within the previous 2 weeks were excluded from the

study. This exclusion was eliminated from the inclusion criteria for enrollment in 1999. Consenting participants donated 2 to 4 ml of venous blood (anticoagulated with EDTA) at the time of enrollment. Three slides per participant, in which each slide contained a thick smear and a thin smear, were prepared from the initial blood specimen. One slide was provided to the clinic personnel for their clinical diagnostic use. Clinic personnel, not study personnel, were responsible for staining and interpreting the results for the slide. The study's investigators retained the other two slides for staining as described previously (10).

Device prototypes. Three different *ParaSight* F+V prototype devices (Becton Dickinson Diagnostic Systems) were evaluated over the course of the 2-year study. In the first year, a single prototype, referred to as FV98 in this report, was available for evaluation with 2,986 patient specimens (Peru, $n = 836$; Thailand, $n = 2,150$). Two additional prototypes were evaluated in the second year and are referred to as FV99-1 and FV99-2 in this report. The FV99-1 prototype was available at the outset of patient enrollment and data collection efforts in 1999 and was evaluated with 1,017 patient specimens (Peru, $n = 351$; Thailand, $n = 666$). The FV99-2 prototype became available later in the 1999 enrollment phase and was tested with a subsequent group of 870 patient specimens (Peru, $n = 393$; Thailand, $n = 477$). Assay procedures and interpretation of the results were performed in strict compliance with the instructions provided by the manufacturer. The ICS methodology used for the *ParaSight* F+V assay involved wicking of a whole-blood lysate onto a nitrocellulose strip and the resulting capture of malarial antigens by antibodies immobilized in the nitrocellulose matrix. Immobilized antibodies were directed against recombinant *P. falciparum* HRP-2- and *P. vivax*-specific antigens. The captured antigens were subsequently visualized by the addition of liposome-encapsulated dye-conjugated secondary ligands. The dye detection system produced a pink test line for bound *P. falciparum* antigen and a blue test line for bound *P. vivax* antigen. The test lines were located approximately 5 mm apart on the assay strip to help with visual recognition and to aid in the interpretation of positive results. Each test strip also contained two internal process-control dotted lines that appeared as positive confirmation of procedure and reagent viabilities. The *ParaSight* F+V test had five possible outcomes on the basis of the presence or absence of test and control lines: negative for *Plasmodium*, positive for *P. falciparum*, positive for *P. vivax*, positive for *P. falciparum* and *P. vivax*, or uninterpretable (i.e., the assay failed to produce a valid positive or negative outcome). Any detectable test line, no matter how faint, was read as a positive result. In addition to being read as a positive result, visible test lines were also compared to a series of photographic standards provided by the manufacturer to determine the relative intensity of the positive test line. Line intensity was scored as 0 (no visible line, a negative test result), 0.25 (the faintest visible line), 0.5, 1, 2, 3, or 4 (the darkest, most intense line color).

Reference microscopy. Technicians serving as study microscopists were prospectively certified as skilled through documentation of training, experience, and past performance (instituted in 1998) and were further required to successfully pass a prestudy microscopy competency examination. For the microscopists to pass this prestudy test, they were required to correctly identify the presence or absence of parasitemia and species for a minimum of 14 of 15 well-characterized Giemsa-stained blood smears (instituted in 1999). One of the two slides of Giemsa-stained blood smears prepared at the time of patient enrollment was independently examined by two study microscopists, as described previously (10). The final diagnostic end point for all positive smears was a calculated level of species-specific parasitemia, expressed as the number of asexual parasites per microliter of whole blood. Each microscopist reviewed 200 oil-immersion high-power fields of a smear before interpreting it as parasite negative. Microscope optics were standardized at both field sites by using microscopes of identical models (model CH-2; Olympus America, Inc., Melville, N.Y.), with identical objectives ($\times 100$ DPlan 100 1.25 oil 160/0.17), and with equivalent eyepieces ($\times 10$ WHK 10 \times /20L or WHK 10 \times /20L-H).

Quality assurance and data integrity. (i) **Blinding of results.** Results from each phase of the study were blinded throughout the data collection and analysis period. Separate teams were organized to conduct patient enrollment, laboratory (MRDD and leukocyte [WBC]) testing, diagnostic microscopy, data entry, and data analysis. In all cases, the results of the *ParaSight* F+V tests were determined prior to the completion of diagnostic microscopy, and the technicians examining the stained smears were strictly blinded to the rapid test results.

(ii) **Microscopy.** In order to ensure a rigorous reference standard, our study team instituted minor changes in the control and validation of microscopy, based on observations from the first year of the study. In 1998, to yield a final microscopic interpretation, the two study microscopists needed to independently agree on three criteria: (i) on the presence or absence of asexual stages of *Plasmodium*; (ii) on the species of *Plasmodium*, when one was present; and (iii) on the calculated level of parasitemia within a factor of 2. The last criterion was refined

by two modifications in 1999. The first change was that results were considered concordant if both study microscopists determined asexual parasitemia levels $\leq 100/\mu\text{l}$, even if the differences between their counts were more than twofold. The second change required that even if counts differed by less than twofold, and so were otherwise concordant, they must also differ by $\leq 50,000/\mu\text{l}$ in absolute terms. The mean for the levels of parasitemia counted by the two microscopists was accepted as the true diagnostic outcome for all slides with concordant results. Results discordant by one or more of these criteria were discarded. Both study slides prepared from that sample were then examined by a third, senior microscopist in a blinded fashion. The cumulative results from both slides, as interpreted by the referee microscopist, were accepted as the true diagnostic outcome. In addition to discordant slide results, the referee microscopists also reviewed 5% of all concordantly interpreted slides for quality assurance purposes (instituted in 1999).

(iii) **WBC determinations.** The Thai study team determined WBC counts using a semiautomated cell counter (Coulter Ac-T 10; Beckman-Coulter, Inc., Fullerton, Calif.). A three-level commercial hematology control material (4C Plus; Beckman-Coulter, Inc.) was analyzed daily, and replicate testing was conducted to monitor the reproducibility between two identical analyzers. At the Iquitos, Peru, test station, a centrifugal hematology system (QBC II; Becton Dickinson Diagnostic Systems) was used to provide WBC counts. The Peruvian study team performed daily instrument calibration checks prior to specimen analysis to ascertain analyzer function within the parameters defined by the manufacturer. Manual WBC counts (UNOPETTE; Becton Dickinson Clinical Laboratory Solutions, Franklin Lakes, N.J.) were performed weekly at both study sites to maintain a backup method for WBC counts as well as to apply an external measure of analyzer function.

(iv) **Assays with prototype MRDDs.** Known positive and negative control materials were supplied by the device manufacturer and were used to verify assay performance on a daily basis. Positive control material consisted of a proprietary formulation known to react at both the *P. falciparum* and *P. vivax* test lines in a manner analogous to that for the targeted malarial antigens. The results for patient specimens were reported as negative, positive for *P. falciparum*, positive for *P. vivax*, positive for *P. falciparum* and *P. vivax*, or uninterpretable, according to the specific instructions provided in the test kits. The assay procedures and interpretation of the results were essentially identical for each of the prototypes. Tests with results reported as uninterpretable, i.e., failure to observe a control line or the presence of a darkly stained background that obscured the test lines, were repeated in an attempt to resolve the discrepant event. Repeatedly uninterpretable results were excluded from calculations of device performance. Investigators and technicians were trained and certified by the study's principal investigator prior to the initiation of testing of patient specimens and operated within a system of daily review and oversight. The *ParaSight F+V* prototype assays evaluated in this study were manufactured in a process compliant with present good manufacturing practices standards established by the FDA to govern the production of in vitro diagnostic devices.

(v) **Protocol review and approval.** The study protocol was reviewed by the Institutional Review Board, Walter Reed Army Institute of Research, and the Human Subjects Safety Review Board, U.S. Army Medical Research and Materiel Command, and were approved as Walter Reed Army Institute of Research Protocol 687 (version 2.21, 1998; version 2.3, 1999). The study protocol was approved annually by the Ministry of Health (MINSa) in Iquitos, Peru, and was performed under the direction of the *Dirección de Salud de Loreto*. Similarly, the study protocol was approved by the Thai Ministry of Public Health for each year of implementation and operated with oversight from Vector-Borne Disease Control Office No. 1, Phrabuddhabat, Thailand.

Data analysis. The "true" diagnostic outcome was defined as the final microscopic interpretation of peripheral blood smears, with microscopy performed as described above. These results were the basis for calculating device sensitivity and specificity. Standard definitions of sensitivity, which was equal to the number of true-positive results/(number of true-positive results + number of false-negative results), and specificity, which was equal to number of true-negative results/(number of true-negative results + number of false-positive results), were used. With regard to MRDD assays interpreted as positive for both *P. falciparum* and *P. vivax* for specimens subsequently found to contain a single species by microscopy, the results were treated as true positive for the species identified by microscopy and false positive for the species not identified by microscopy. Device sensitivity was also determined for each of several prospectively defined ranges of parasitemia: >0 to 500, 501 to 1,000, 1,001 to 5,000, and $>5,000$ parasites/ μl . Receiver operating characteristic (ROC) curves were constructed and analyzed to determine the relative accuracy of each prototype at diagnostic thresholds corresponding to test line intensity for positive MRDD assay results (13, 43, 44). ROC curves plot the fraction of false-positive results of each prototype (1 minus

the specificity; the ability to distinguish true-negative results from false-positive results) against the fraction of true-positive results (sensitivity; the ability to distinguish true-positive results from false-negative results) for each of the six possible positive line intensity values (0.25 through 4). This analysis displayed trade-offs between sensitivity and specificity that occurred as the result of incremental changes in positive test line intensity. The overall accuracy of each device was quantified with an area under the curve (AUC) value with the standard error and an estimate of its 95% confidence interval (CI_{95}) (9, 38). Each prototype assay's performance for the detection of single species was further evaluated by comparison of the optimal diagnostic cutoff values. The optimal cutoff point was represented as the test line intensity with the maximum sum of sensitivity and specificity for each prototype. Diagnostic accuracy was interpreted by considering a combination of greater AUC values and lower diagnostic thresholds, i.e., the optimized relationship between test line intensity and a correct diagnostic result. Documentation was maintained to ensure that all study data, interpretations, and calculated values were traceable to original source material. Statistical significance was set at a *P* value <0.05 , and AccuROC (version 2.5) software (Accumetric Corporation, Montreal, Quebec, Canada) was used to analyze ROC curve data.

RESULTS

Study demographics. A total of 4,894 patients were enrolled during the 2-year study period: 1,588 (32.4%) and 3,306 (67.6%) from Peru and Thailand, respectively. Among these patients, 14 (0.3%) patients (13 patients in 1998 and 1 patient in 1999) were eventually excluded from the study group. Among the remaining 4,880 evaluable patients who provided blood specimens for assay, assay results for a total of 7 (0.1%) patients were found to be invalid for the cumulative enrollment population, primarily as a result of incomplete documentation. Despite the eligibility of younger patients, demographic data collected for the volunteers showed that no subjects younger than 18 years of age had enrolled at either the Thai or the Peruvian site during the 1999 season. The data generated for the remaining 4,873 specimens were included in the study data set.

Diagnostic microscopy. Microscopic examination of 4,873 peripheral blood smears showed that a total of 2,051 (42.1%) were positive for malaria, with the remaining 2,822 (57.9%) determined to be negative for asexual parasite stages (Table 1). Species representation varied slightly between enrollment years, with the largest variation noted in the number of *P. falciparum*-positive blood smears detected in the Peruvian cohort: 115 (13.8%) in 1998 and 63 (8.5%) in 1999. The only *P. malariae*-positive blood smears were found in the Thai cohort, and both of these were detected in participants enrolled in 1998. A total of 50 (1.0%) mixed infections with both *P. falciparum* and *P. vivax* were detected and were found in 31 (1.0%) specimens obtained in 1998 and 19 (1.0%) specimens obtained in 1999. Study microscopists examining blood smears reported interpretations concordant for all three microscopic criteria (presence of parasitemia, species, and parasite density within twofold) for 93.8% of smears in 1998 and for 95.7% of smears in 1999.

MRDD assay performance for *P. falciparum*. A cross-tabulation of the *ParaSight F+V* test results for *P. falciparum* against the reference microscopy results is shown in Table 2. The greatest proportion of false-negative results for *P. falciparum* was observed at the lowest parasite concentrations with all three of the prototypes. With regard to the false-positive results for *P. falciparum*, a relatively large number of FV98 assays produced very faint test lines (line intensity, 0.25) for both *P.*

TABLE 1. Summary of microscopic findings

Result ^a	No. (%) of specimens positive				Total
	1998		1999		
	Peru	Thailand	Peru	Thailand	
<i>P. falciparum</i>	115 (13.8)	431 (20.0)	63 (8.5)	250 (21.9)	859 (17.6)
<i>P. vivax</i>	207 (24.8)	450 (20.9)	222 (29.8)	261 (22.8)	1,140 (23.4)
Mixed <i>P. falciparum</i> and <i>P. vivax</i>	12 (1.4)	19 (0.9)	10 (1.3)	9 (0.8)	50 (1.0)
<i>P. malariae</i>	0 (0)	2 (0.09)	0 (0)	0 (0)	2 (0.04)
Negative ^b	502 (60.0)	1,248 (58.0)	449 (60.3)	623 (54.5)	2,822 (57.9)
Study total	836 (100)	2,150 (100)	744 (100)	1,143 (100)	4,873 (100)

^a Positive microscopy results required the presence of asexual *Plasmodium* stages.

^b Blood smears containing *Plasmodium* sp. gametocytes in the absence of asexual stages of the same species were interpreted as negative.

falciparum and *P. vivax*. The appearance of faint test lines affected the performance of FV98, with 2,113 (71.0%) of all specimens interpreted, according to the manufacturer’s instructions, as having mixed *P. falciparum* and *P. vivax* infections, given the presence of test lines for both *P. falciparum* and *P. vivax*. Among 1,829 false-positive results for *P. falciparum* with the FV98 prototype, 940 (51.4%) of the positive test lines were interpreted to have a line intensity of 0.25. The number of false-positive lines for *P. falciparum* at the 0.25 intensity level was significantly reduced with the 1999 prototypes, particularly FV99-2.

Given the lack of an available method for the precise determination of the levels of parasitemia for individual species in true mixed infections and the overall small number of mixed infections in the total study population (1.0%), device sensitivity and specificity were calculated by exclusion of the results for mixed infections defined microscopically. As noted, uninterpretable results, i.e., the results of repeat tests that failed to

produce an interpretable diagnostic outcome, were also excluded from the final data set used to determine device performance. The largest number of uninterpretable results was observed with the FV98 prototype, with 64 of 2,986 specimens (2.1%) failing to produce a clear diagnostic result after repeat testing.

The sensitivity of the FV98 prototype for detection of *P. falciparum* relative to the reference microscopy results was 98% (CI₉₅, 97 to 100%) (Table 3). The FV99-1 and FV99-2 assays demonstrated overall sensitivities of 94% (CI₉₅, 89 to 97%) and 98% (CI₉₅, 94 to 100%), respectively. Both the FV99-1 and FV99-2 assays were 100% sensitive in detecting *P. falciparum* when the level of parasitemia exceeded 500/μl. There were no significant differences in device sensitivities between the Peruvian and Thai cohorts for any of the parasitemia ranges examined or among any of the three prototypes. Each of the three assays showed progressively decreasing sensitivities as the parasite concentrations decreased. The speci-

TABLE 2. Cross-tabulation of *ParaSight* F+V results for *P. falciparum* against microscopy results^a

Microscopy result	No. of specimens with the indicated result with the following prototype:								
	FV98			FV99-1			FV99-2		
	Negative ^b	Positive ^c	Mixed	Negative ^b	Positive ^c	Mixed	Negative ^b	Positive ^c	Mixed
Positive for <i>P. falciparum</i> (no. of parasites/μl)									
>0–500	2 (2/0)	25	52	8 (6/2)	13	12	4 (4/0)	19	1
501–1,000	2 (2/0)	9	20	0 (0/0)	9	1	0 (0/0)	7	1
1,001–5,000	1 (1/0)	47	54	0 (0/0)	19	3	0 (0/0)	31	1
>5,000	4 (4/0)	110	216	1 (1/0)	58	25	0 (0/0)	77	23
Total	9 (9/0)	191	342	9 (7/0)	99	41	4 (4/0)	134	26
Negative for <i>P. falciparum</i>									
Negative for <i>Plasmodium</i>	317 (304/13)	68	1,313	491 (358/133)	11	110	436 (394/42)	10	13
Positive for <i>P. vivax</i>	202 (36/166)	12	435	225 (21/204)	3	21	211 (29/182)	2	21
Positive for <i>P. malariae</i>	1 (1/0)	1	0	0 (0/0)	0	0	0 (0/0)	0 (0/0)	0 (0/0)
Total	520 (341/179)	81	1,748	716 (379/337)	14	131	647 (423/224)	12	34
Mixed infection									
<i>P. falciparum</i> + <i>P. vivax</i>	2 (1/1)	6	23	4 (0/4)	1	2	3 (2/1)	5	4
Total	2 (1/1)	6	23	4 (0/4)	1	2	3 (2/1)	5	4
Study total	531 (351/180)	278	2,113	729 (386/343)	114	174	654 (429/225)	151	64

^a Excludes uninterpretable MRDD assay results (FV98, n = 64; FV99-1, n = 0; FV99-2, n = 1).

^b The values in parentheses are the number of specimens with negative MRDD assay results/number of specimens positive only for *P. vivax*, i.e., negative results for *P. falciparum*.

^c MRDD assay results positive only for *P. falciparum*.

TABLE 3. Performance characteristics of ParaSight F+V prototypes for detection of *P. falciparum*^a

Performance parameter	FV98				FV99-1				FV99-2			
	No. of specimens		% of specimens (CI ₉₅)		No. of specimens		% of specimens (CI ₉₅)		No. of specimens		% of specimens (CI ₉₅)	
	Peru	Thailand	Overall		Peru	Thailand	Overall		Peru	Thailand	Overall	
Sensitivity	542	95 (89-98)	99 (98-100)	98 (97-100)	49	92 (75-100)	94 (89-98)	94 (89-97)	164	95 (82-100)	98 (94-100)	98 (94-100)
Any level of parasitemia >0/μl	79	100 (74-100)	97 (90-100)	97 (91-100)	33	83 (36-100)	74 (54-89)	76 (58-89)	24	67 (22-96)	89 (65-99)	83 (63-95)
>0-500/μl	31	80 (44-97)	100 (84-100)	94 (79-100)	10	100 (3-100)	100 (66-100)	100 (69-100)	8	100 (3-100)	100 (59-100)	100 (63-100)
501-1,000/μl	102	100 (86-100)	99 (93-100)	99 (95-100)	22	100 (40-100)	100 (81-100)	100 (85-100)	32	100 (74-100)	100 (83-100)	100 (89-100)
1,001-5,000/μl	330	94 (86-98)	100 (99-100)	99 (97-100)	84	93 (68-100)	100 (95-100)	99 (94-99)	100	100 (81-100)	100 (96-100)	100 (96-100)
>5,000/μl												
Specificity	2,349	35 (31-38)	17 (15-19)	22 (20-24)	86	83 (79-87)	83 (80-86)	83 (80-86)	694	97 (95-99)	89 (85-92)	93 (91-95)

^a Mixed infections (*P. falciparum* and *P. vivax*) detected by microscopy were excluded from calculations of device performance (1998, *n* = 31; 1999, *n* = 19).

ficity of FV98 for exclusion of the presence of *P. falciparum* was 22% (CI₉₅, 20 to 24%), with values of 83% (CI₉₅, 80 to 86%) and 93% (CI₉₅, 91 to 95%) for FV99-1 and FV99-2, respectively. The specificities of the FV98 prototype differed significantly at the Peru and Thai study sites.

MRDD assay performance for *P. vivax*. A cross-tabulation of ParaSight F+V test results for *P. vivax* against the reference microscopy results is shown in Table 4. For all three prototypes, most of the false-negative results (75 of 103 [72.8%]) were demonstrated with levels of parasitemia in the range of >0 to 500 parasites/μl, consistent with the observations reported for assay performance for *P. falciparum*. Two *P. vivax*-positive samples that had levels of parasitemia >5,000 parasites/μl tested negative with the FV98 prototype. With the FV99-2 prototype, one sample with a level of *P. vivax* parasitemia >5,000/μl tested negative for *P. vivax* but had a positive result for *P. falciparum*. As noted for the performance of the assays for *P. falciparum*, the appearance of faint positive lines (line intensity, 0.25) detracted from the performance of the FV98 prototype, with 435 of 649 (67%) specimens positive for *P. vivax* by reference microscopy interpreted as containing a mixture of *P. falciparum* and *P. vivax* and 1,655 of 2,242 (74%) of *P. vivax*-negative samples interpreted as containing a mixture of *P. falciparum* and *P. vivax*. Among a total of 1,668 false-positive results for *P. vivax* obtained with the FV98 prototype, 915 (54.9%) had test lines positive for *P. vivax* with an intensity of 0.25. Among specimens found to be negative for any *Plasmodium* sp. microscopically, the proportions of false-positive assay results for *P. vivax* were 78.1, 39.7, and 12.0% for prototypes FV98, FV99-1, and FV99-2, respectively.

The sensitivity of the FV98 prototype for the detection of *P. vivax* relative to the reference microscopy results was 93% (CI₉₅, 90 to 94%) (Table 5). The FV99-1 and FV99-2 assays demonstrated overall sensitivities for the detection of *P. vivax* of 90% (CI₉₅, 86 to 94%) and 87% (CI₉₅, 82 to 91%), respectively. There were no significant differences in device sensitivities between the Peruvian and Thai study sites for any of the parasitemia ranges. All three prototypes showed decreased sensitivities for the detection of *P. vivax* with lower levels of parasitemia, most notably, levels <500 parasites/μl. The overall specificities for exclusion of the presence of *P. vivax* in specimens with true-negative results were 26% (CI₉₅, 24 to 27%) for FV98, 62% (CI₉₅, 59 to 66%) for FV99-1, and 87% (CI₉₅, 84 to 89%) for FV99-2. The differences in specificities for *P. vivax* between study sites were statistically significant for all three prototypes. FV98 and FV99-2 showed greater specificities in Peru, and FV99-1 showed a greater specificity in Thailand.

Diagnostic accuracy. Figure 1 shows ROC curves for each of the ParaSight F+V prototypes when test line intensity was used as the decision cutoff point for diagnostic outcome; curves were constructed for the performance of the assays for the detection of both *P. falciparum* (Fig. 1A) and *P. vivax* (Fig. 1B). The AUC for the diagnostic performance of FV98 for the detection of *P. falciparum* was significantly different from the AUCs for the performances of both the FV99-1 and the FV99-2 prototypes for the detection of *P. falciparum* (Table 6). The AUC for the diagnostic performance of FV99-2 for the detection of *P. falciparum* was significantly different from the AUC for the performance of the FV99-1 prototype for the

TABLE 4. Cross-tabulation of *ParaSight* F+V results for *P. vivax* against microscopy results^a

Microscopy result	No. of specimens with the indicated result with the following prototype:								
	FV98			FV99-1			FV99-2		
	Negative ^b	Positive ^c	Mixed	Negative ^b	Positive ^c	Mixed	Negative ^b	Positive ^c	Mixed
Positive for <i>P. vivax</i>									
(no. for parasites/μl)									
>0-500	31 (24/7)	2	119	20 (18/2)	30	9	24 (24/0)	24	5
501-1,000	7 (6/1)	6	31	1 (1/0)	17	1	1 (1/0)	14	1
1,001-5,000	8 (5/3)	46	127	3 (2/1)	59	5	5 (4/1)	56	3
>5,000	2 (1/1)	112	158	0 (0/0)	98	6	1 (0/1)	88	12
Total	48 (36/12)	166	435	24 (21/3)	204	21	31 (29/2)	182	21
Negative for <i>P. vivax</i>									
Negative for <i>Plasmodium</i>	372 (304/68)	13	1,313	369 (358/11)	133	110	404 (394/10)	42	13
Positive for <i>P. falciparum</i>	200 (9/191)	0	342	100 (7/99)	2	41	138 (4/134)	0	26
Positive for <i>P. malariae</i>	2 (1/1)	0	0	0 (0/0)	0	0	0 (0/0)	0	0
Total	574 (314/260)	13	1,655	475 (365/110)	135	151	542 (398/144)	42	39
Mixed infection									
<i>P. falciparum</i> + <i>P. vivax</i>	7 (1/6)	1	23	1 (0/1)	4	2	7 (2/5)	1	4
Total	7 (1/6)	1	23	1 (0/1)	4	2	7 (2/5)	1	4
Study total	629 (351/278)	180	2,113	500 (386/114)	343	174	580 (429/151)	225	64

^a Excludes uninterpretable MRDD assay results (FV98, n = 64; FV99-1, n = 0; FV99-2, n = 1).

^b The values in parentheses are the number of specimens with negative MRDD assay results/number of specimens positive only for *P. falciparum*, i.e., negative results for *P. vivax*.

^c MRDD assay results positive only for *P. vivax*.

detection of *P. falciparum*. There were no statistically significant differences in AUCs for the performances of the three prototypes for the detection of *P. vivax*. The optimal diagnostic cutoff for each MRDD assay was determined by maximizing the sum of the sensitivity and the specificity determined for each discrete threshold of positive line intensity. For diagnostic accuracy for the detection of *P. falciparum*, optimal cutoffs were determined to be at line intensities of 1 for FV98, 1 for FV99-1, and 0.5 for FV99-2. For diagnostic accuracy for the detection of *P. vivax*, optimal cutoffs were determined to be at line intensities of 1 for FV98, 1 for FV99-1, and 0.25 for FV99-2.

DISCUSSION

The present study determined the performance characteristics of three prototype *ParaSight* F+V assays, all of which were engineered to detect two malaria species in a single test strip. The study protocol was specifically designed to evaluate the performances of the prototypes with specimens from a self-presenting symptomatic patient population in areas where malaria is endemic. The findings reported here are for 2 years of a multiyear program being conducted to evaluate MRDD assays of various formats and from multiple manufacturing sources. The program evolved from the investigation of an assay device for the detection of a single species, the *ParaSight* F assay, to the present study with the specific intent of identifying one or more devices capable of satisfying prospectively defined performance criteria. The clinical capability envisioned for usage in the field with an immunologically naïve population necessitates reliability at high levels of parasitemia, a lower threshold of detection of at least 500 parasites/μl, elimination of the need for cold-chain conditions for assay transport and

storage, and the detection of multiple species, with an emphasis, in order of priority, on *P. falciparum* and *P. vivax*. It was further required that field assays be uncomplicated in their operation and interpretation and, ultimately, suitable for regulatory review and approval by the FDA.

The findings generated over the course of 2 years show the progressive refinement of the *ParaSight* F+V assay to a point, exemplified by prototype FV99-2, that it provided an accurate means of detection of both *P. falciparum* and *P. vivax* in two distinct study cohorts, one in Asia and the other in South America. Compared to the reference microscopy results, the FV99-2 prototype device correctly identified asexual parasites in 98% of all *P. falciparum*-positive peripheral blood specimens and 87% of all *P. vivax*-positive peripheral blood specimens. These sensitivity values were accompanied by specificities of 93 and 87% for *P. falciparum* and *P. vivax*, respectively. Of further note, the FV99-2 device achieved 100% sensitivity with peripheral blood samples (n = 140) containing *P. falciparum* parasites at >500/μl. Such highly reliable sensitivities at a relatively low level of parasitemia suggest that, in patients suspected of having malaria, if the result of the initial test is negative, perhaps due to a very low level of parasitemia, testing of serially collected blood samples, perhaps at 12- to 24-h intervals, would yield very accurate sensitivities. Such a testing strategy would mirror that already used by competent microscopists for serial examination of blood smears (14). This level of diagnostic performance exceeds that of clinical microscopy reported in a variety of settings and has the potential to significantly enhance diagnostic capabilities in those instances in which skilled microscopy is not readily available (1, 6, 21, 23, 26).

It is not entirely clear why performance differences were observed in some instances between the two study sites. Con-

TABLE 5. Performance characteristics of *ParaSight* F+V prototypes for detection of *P. vivax*^a

Performance parameter	FV98 %			FV99-1 %			FV99-2 %					
	% of specimens (CI ₉₅)		No. of specimens	% of specimens (CI ₉₅)		No. of specimens	% of specimens (CI ₉₅)		No. of specimens			
	Peru	Thailand		Overall	Peru		Thailand	Overall		Peru	Thailand	Overall
Sensitivity												
Any level of parasitemia >0/μl	649	93 (88–96)	93 (90–95)	249	97 (91–100)	86 (79–91)	90 (86–94)	234	85 (78–91)	88 (81–94)	87 (82–91)	
>0–500/μl	152	67 (43–85)	82 (74–88)	59	83 (52–98)	62 (46–75)	66 (53–78)	53	46 (27–67)	63 (42–81)	55 (40–68)	
501–1,000/μl	44	87 (60–98)	83 (64–94)	19	88 (47–100)	100 (72–100)	95 (74–100)	16	88 (47–100)	100 (63–100)	94 (70–100)	
1,001–5,000/μl	181	92 (82–97)	98 (93–100)	67	100 (89–100)	92 (78–98)	96 (87–100)	64	92 (78–98)	93 (76–100)	92 (83–97)	
>5,000/μl	272	99 (95–100)	99 (97–100)	104	100 (93–100)	100 (94–100)	100 (97–100)	101	100 (93–100)	98 (89–100)	99 (95–100)	
Specificity	2,242	39 (35–43)	21 (19–23)	76	44 (37–50)	72 (67–75)	62 (59–66)	623	94 (90–96)	82 (77–86)	87 (84–89)	

^a Mixed infections (*P. falciparum* and *P. vivax*) detected by microscopy were excluded from calculations of device performance (1998, *n* = 31; 1999, *n* = 19).

siderable rigor was exerted in standardizing the execution of the protocol between the two study sites and in validating the performance parameters. Confounding variables, such as tester-dependent bias or subtle differences in sample handling, may have played a role in the variation in the results obtained between the Thai and Peruvian study sites.

The *P. falciparum* portion of the *ParaSight* F+V prototypes was based on a predicate device, the *ParaSight* F test, which was manufactured and distributed for several years prior to the development of the *ParaSight* F+V format. As reported previously, the *ParaSight* F device demonstrated a high level of performance under study conditions similar to those used in the present study (10). The development of a multispecies prototype assay based on the predicate assay for a single species likely required refinements in materials, device construction, and manufacturing processes and optimization of reagent formulations. While these refinements are the proprietary property of the developer, it is clear that the incorporation of additional test lines into the assay matrix affected assay performance during the progressive development of the device. The presence of additional antibody-impregnated test and control lines on the nitrocellulose dipstick and the resulting impacts on sample flow and secondary ligand binding were factors in the observed performances of early prototypes. The precise modifications to the FV99-2 prototype that alleviated the deleterious effects of false-positive results because of low line intensities are not known to investigators in the field, but the findings of the present study demonstrate a marked improvement in the performance of this final prototype device.

It was difficult to reliably assess the ability of the *ParaSight* F+V prototypes to accurately discriminate true mixed infections. The lack of a precise method to determine the contribution of each species to the total parasitemia in a mixed infection and the small number of microscopically detected mixed infections in each cohort prompted the exclusion of mixed infections from calculations of device performance. A better reference method for identifying mixed infections and discriminating the proportional contribution of each individual species to the total parasitemia is necessary to fully assess MRDD assay performance in the detection of mixed infections. This is an area of investigation that may benefit from the use of a molecular diagnostic technology to augment skilled microscopy as a reference standard (24, 29–31, 35, 41, 42).

The data presented in Tables 2 and 4 demonstrate an increase in device sensitivities, i.e., the ability to detect *Plasmodium* when it was present, with increased parasite concentrations. It is not clear why a small number of specimens with parasitemias >5,000 parasites/μl tested falsely negative, although this observation is not without precedent in either HRP-2-based assays (4, 10, 15, 19, 22, 28, 39) or MRDD assays that exploit other malarial antigens (18, 19, 27). While the number of specimens with false-negative results and high levels of parasitemia was small (<1.0% for each prototype), it is an area of continuing concern and suggests the interaction of other factors that may influence test outcome for a small subpopulation of symptomatic patients.

ROC analysis provides insight into the performances of the prototypes by illustrating the relationship between line intensity and diagnostic accuracy. The AUCs calculated from the ROC curves in Fig. 1 are indicative of the accuracies of the

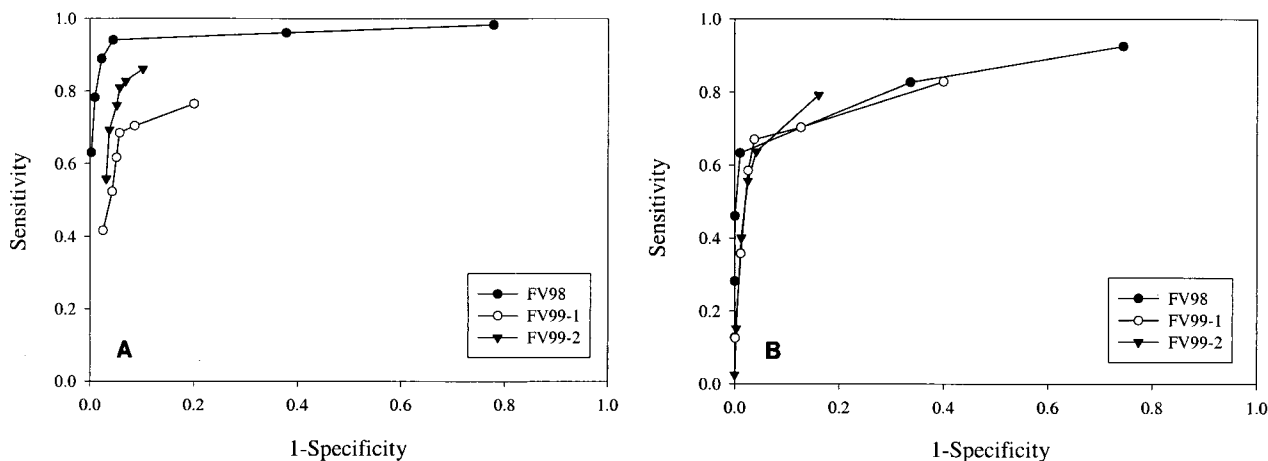


FIG. 1. ROC curves for diagnostic performance of *ParaSight* F+V prototype assays. (A) Prototype performance for detection of *P. falciparum*. (B) Prototype performance for detection of *P. vivax*.

devices across the full range of potential test line intensities. Initial assessment of AUCs suggested that the FV98 MRDD was the most accurate device for the detection of *P. falciparum* in the enrolled patient population. Closer scrutiny of maximized diagnostic cutoffs revealed that the FV98 assay was optimized at a line intensity of 1 and, consequently, experienced significant performance degradation at the 0.25 and 0.5 line intensities. The FV99-1 prototype experienced a similar pattern of performance optimized at an intensity cutoff of 1. Although the AUC for the FV99-2 prototype was lower than that for the FV98 prototype, the FV99-2 prototype achieved an optimal cutoff at a line intensity of 0.5, indicating an improvement in diagnostic performance in association with lower-intensity test lines. A similar observation was made with respect to test accuracy for the detection of *P. vivax*, with the FV99-2 assay exhibiting optimized performance at a line intensity cutoff of 0.25, the faintest line visible to test readers. An MRDD test system that can be accurately and reliably interpreted as positive with the presence of any visible test line, regardless of intensity, enjoys a decided advantage in a point-of-care diagnostic setting. Conversely, an assay that offers an equivocal outcome at low line intensities will erode end-user confidence. The analysis of optimized line intensity is not intended to

suggest that rapid diagnostic assays be interpreted on the basis of line intensity. Rather, these data illustrate the trade-offs between sensitivity and specificity and demonstrate the manufacturer's success in optimizing performance across the full spectrum of assay reactivities.

The present study documents the performance of three distinct prototypes resulting from the iterative development and refinement of the *ParaSight* F+V assay. We have demonstrated the benefits of using large-scale, multisite field trials to assess product refinement and for product evaluation. This investigation facilitated a critical assessment of the diagnostic performances of multiple generations of prototypes with specimens from a large population of symptomatic patients. The findings presented herein identify the *ParaSight* F+V test, in the FV99-2 format, as a mature MRDD assay capable of clinically acceptable diagnostic performance in a field setting. Notwithstanding the performances of the prototypes evaluated in this study, the manufacturer has no plans at present to complete the development and commercialization of the *ParaSight* F+V MRDD test.

ACKNOWLEDGMENTS

We thank Joe Perrone and David Baggett of Becton Dickinson Diagnostic Systems for providing access to the *ParaSight* F+V prototype test kits used in this study. We also acknowledge Janelle Rhorer of Statistics Collaborative, Washington, D.C., for statistical analysis of our data. Finally, we express our appreciation for the skill and dedication of the field study teams deployed by the Navy Medical Research Center, Lima, Peru, and the Armed Forces Research Institute for Medical Sciences, Bangkok, Thailand.

REFERENCES

1. Aron, J. L. 1982. Malaria epidemiology and detectability. *Trans. R. Soc. Med. Hyg.* 76:595-601.
2. Banchongaksorn, T., S. Prajakwong, W. Rooney, and P. Vickers. 1997. Operational trial of *ParaSight*-F (dipstick) in the diagnosis of falciparum malaria at the primary health care level. *Southeast Asian J. Trop. Med. Public Health* 28:243-246.
3. Barat, L., J. Chipipa, M. Kolczak, and T. Sukwa. 1999. Does the availability of blood slide microscopy for malaria at health centers improve the management of persons with fever in Zambia? *Am. J. Trop. Med. Hyg.* 60:1024-1030.
4. Beadle, C., G. W. Long, W. R. Weiss, P. D. McElroy, S. M. Maret, A. J. Oloo, and S. L. Hoffman. 1994. Diagnosis of malaria by detection of Plasmodium

TABLE 6. Diagnostic accuracy of *ParaSight* F+V prototypes determined by ROC analysis

Prototype and organism	ROC AUC ^a	CI ₉₅	SE
FV98			
<i>P. falciparum</i>	0.9630	0.9509-0.9751	0.0062
<i>P. vivax</i>	0.8473	0.8269-0.8677	0.0104
FV99-1			
<i>P. falciparum</i>	0.8318	0.7920-0.8717	0.0203
<i>P. vivax</i>	0.8359	0.8027-0.8692	0.0169
FV99-2			
<i>P. falciparum</i>	0.8907	0.8598-0.9217	0.0158
<i>P. vivax</i>	0.8522	0.8218-0.8826	0.0155

^a AUCs were determined from analysis of the ROC plots shown in Fig. 1.

- falciparum HRP-2 antigen with a rapid dipstick antigen-capture assay. *Lancet* **343**:564–568.
5. Caraballo, A., and A. Ache. 1996. The evaluation of a dipstick test for *Plasmodium falciparum* in mining areas of Venezuela. *Am. J. Trop. Med. Hyg.* **55**:482–484.
 6. Collier, J. A., and J. M. Longmore. 1983. The reliability of the microscopic diagnosis of malaria in the field and in the laboratory. *Ann. Trop. Med. Parasitol.* **77**:113–117.
 7. Craig, M. H., and B. L. Sharp. 1997. Comparative evaluation of four techniques for the diagnosis of *Plasmodium falciparum* infections. *Trans. R. Soc. Trop. Med. Hyg.* **91**:279–282.
 8. Cropley, I. M., D. N. Lockwood, D. Mack, G. Pasvol, and R. N. Davidson. 2000. Rapid diagnosis of falciparum malaria by using the ParaSight F test in travellers returning to the United Kingdom: prospective study. *BMJ* **321**:484–485.
 9. DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**:837–845.
 10. Forney, J. R., A. J. Magill, C. Wongsrichanalai, J. Sirichaisinthop, C. T. Bautista, D. G. Heppner, R. S. Miller, C. F. Ockenhouse, A. Gubanov, R. Shafer, C. C. DeWitt, H. A. Quino-Ascurra, K. E. Kester, K. C. Kain, D. S. Walsh, W. R. Ballou, and R. A. Gasser, Jr. 2001. Malaria rapid diagnostic devices: performance characteristics of the ParaSight F device determined in a multisite field study. *J. Clin. Microbiol.* **39**:2884–2890.
 11. Gautam, A. S., R. C. Sharma, R. M. Bhatt, and D. K. Gupta. 1992. Microscopic diagnosis of malaria in Kheda District of Gujarat. *Indian J. Malariol.* **29**:83–87.
 12. Gaye, O., M. Diouf, E. F. Dansokho, G. McLaughlin, and S. Diallo. 1998. Diagnosis of *Plasmodium falciparum* malaria using ParaSight F, ICT malaria PF and malaria IgG CELISA assays. *Parasite* **5**:189–192.
 13. Henderson, A. R. 1993. Assessing test accuracy and its clinical consequences: a primer for receiver operating characteristic curve analysis. *Ann. Clin. Biochem.* **30**:521–539.
 14. Hommel, M. 2002. Diagnostic methods in malaria, p. 35–58. *In* D. A. Warrel and H. M. Gilles (ed.), *Essential malariology*, 4th ed. Arnold, London, United Kingdom.
 15. Humar, A., C. Ohrt, M. A. Harrington, D. Pillai, and K. C. Kain. 1997. Parasight F test compared with the polymerase chain reaction and microscopy for the diagnosis of *Plasmodium falciparum* malaria in travelers. *Am. J. Trop. Med. Hyg.* **56**:44–48.
 16. Iqbal, J., P. R. Hira, A. Sher, and A. A. Al-Enezi. 2001. Diagnosis of imported malaria by *Plasmodium* lactate dehydrogenase (pLDH) and histidine-rich protein 2 (PfHRP-2)-based immunocapture assays. *Am. J. Trop. Med. Hyg.* **64**:20–23.
 17. Iqbal, J., A. Sher, P. R. Hira, and R. Al-Owaish. 1999. Comparison of the OptiMAL test with PCR for diagnosis of malaria in immigrants. *J. Clin. Microbiol.* **37**:3644–3646.
 18. Jelinek, T., M. P. Grobusch, and H. D. Nothdurft. 2000. Use of dipstick tests for the rapid diagnosis of malaria in nonimmune travelers. *J. Travel Med.* **7**:175–179.
 19. Jelinek, T., M. P. Grobusch, S. Schwenke, S. Steidl, F. von Sonnenburg, H. D. Nothdurft, E. Klein, and T. Loscher. 1999. Sensitivity and specificity of dipstick tests for rapid diagnosis of malaria in nonimmune travelers. *J. Clin. Microbiol.* **37**:721–723.
 20. Kachur, S. P., E. Nicolas, V. Jean-Francois, A. Benitez, P. B. Bloland, Y. Saint Jean, D. L. Mount, T. K. Ruebush II, and P. Nguyen-Dinh. 1998. Prevalence of malaria parasitemia and accuracy of microscopic diagnosis in Haiti, October 1995. *Rev. Panam. Salud Publica* **3**:35–39.
 21. Kain, K. C., M. A. Harrington, S. Tennyson, and J. S. Keystone. 1998. Imported malaria: prospective analysis of problems in diagnosis and management. *Clin. Infect. Dis.* **27**:142–149.
 22. Kilian, A. H., G. Kabagambe, W. Byamukama, P. Langi, P. Weis, and F. von Sonnenburg. 1999. Application of the ParaSight-F dipstick test for malaria diagnosis in a district control program. *Acta Trop.* **72**:281–293.
 23. Kilian, A. H., W. G. Metzger, E. J. Mutschelknauss, G. Kabagambe, P. Langi, R. Korte, and F. von Sonnenburg. 2000. Reliability of malaria microscopy in epidemiological studies: results of quality control. *Trop. Med. Int. Health* **5**:3–8.
 24. Laoboonchai, A., F. Kawamoto, N. Thanosingha, S. Kojima, R. R. S. Miller, K. C. Kain, and C. Wongsrichanalai. 2001. PCR-based ELISA technique for malaria diagnosis of specimens from Thailand. *Trop. Med. Int. Health* **6**:458–462.
 25. Lema, O. E., J. Y. Carter, N. Nagelkerke, M. W. Wangai, P. Kitenge, S. M. Gikunda, P. A. Arube, C. G. Munafu, S. F. Materu, C. A. Adhiambo, and H. K. Mukunza. 1999. Comparison of five methods of malaria detection in the outpatient setting. *Am. J. Trop. Med. Hyg.* **60**:177–182.
 26. Milne, L. M., M. S. Kyi, P. L. Chiodini, and D. C. Warhurst. 1994. Accuracy of routine laboratory diagnosis of malaria in the United Kingdom. *J. Clin. Pathol.* **47**:740–742.
 27. Palmer, C. J., J. F. Lindo, W. I. Klaskala, J. A. Quesada, R. Kaminsky, M. K. Baum, and A. L. Ager. 1998. Evaluation of the OptiMAL test for rapid diagnosis of *Plasmodium vivax* and *Plasmodium falciparum* malaria. *J. Clin. Microbiol.* **36**:203–206.
 28. Pieroni, P., C. D. Mills, C. Ohrt, M. A. Harrington, and K. C. Kain. 1998. Comparison of the ParaSight-F test and the ICT Malaria Pf test with the polymerase chain reaction for the diagnosis of *Plasmodium falciparum* malaria in travellers. *Trans. R. Soc. Trop. Med. Hyg.* **92**:166–169.
 29. Postigo, M., A. Mendoza-Leon, and H. A. Perez. 1998. Malaria diagnosis by the polymerase chain reaction: a field study in south-eastern Venezuela. *Trans. R. Soc. Trop. Med. Hyg.* **92**:509–511.
 30. Relf, W. A., R. E. Boreham, P. Tapchaisri, S. Khusmith, A. Healey, P. Upcroft, S. Tharavanij, and C. Kidson. 1990. Diagnosis of *Plasmodium vivax* malaria using a specific deoxyribonucleic acid probe. *Trans. R. Soc. Trop. Med. Hyg.* **84**:630–634.
 31. Rubio, J. M., J. Roche, P. J. Berzosa, E. Moyano, and A. Benito. 2000. The potential utility of the semi-nested multiplex PCR technique for the diagnosis and investigation of congenital malaria. *Diagn. Microbiol. Infect. Dis.* **38**:233–236.
 32. Shiff, C. J., Z. Premji, and J. N. Minjas. 1993. The rapid manual Parasight® F test. A new diagnostic tool for *Plasmodium falciparum* infection. *Trans. R. Soc. Trop. Med. Hyg.* **87**:646–648.
 33. Singh, N., A. Saxena, and N. Valecha. 2000. Field evaluation of the ICT malaria P.f/P.v immunochromatographic test for diagnosis of *Plasmodium falciparum* and *P. vivax* infection in forest villages of Chhindwara, central India. *Trop. Med. Int. Health* **5**:765–770.
 34. Singh, N., N. Valecha, and V. P. Sharma. 1997. Malaria diagnosis by field workers using an immunochromatographic test. *Trans. R. Soc. Trop. Med. Hyg.* **91**:396–397.
 35. Tham, J. M., S. H. Lee, T. M. Tan, R. C. Ting, and U. A. Kara. 1999. Detection and species determination of malaria parasites by PCR: comparison with microscopy and with ParaSight-F and ICT malaria Pf tests in a clinical environment. *J. Clin. Microbiol.* **37**:1269–1273.
 36. Thomson, S., R. C. Lohmann, L. Crawford, R. Dubash, and H. Richardson. 2000. External quality assessment in the examination of blood films for malarial parasites within Ontario, Canada. *Arch. Pathol. Lab. Med.* **124**:57–60.
 37. Tjitra, E., S. Suprianto, M. Dyer, B. J. Currie, and N. M. Anstey. 1999. Field evaluation of the ICT malaria P.f/P.v immunochromatographic test for detection of *Plasmodium falciparum* and *Plasmodium vivax* in patients with a presumptive clinical diagnosis of malaria in eastern Indonesia. *J. Clin. Microbiol.* **37**:2412–2417.
 38. Vida, S. 1993. A computer program for non-parametric receiver operating characteristic analysis. *Comput. Methods Programs Biomed.* **40**:95–101.
 39. Wongsrichanalai, C., N. Chuanak, S. Tulyayon, N. Thanosingha, A. Laoboonchai, K. Thimasarn, T. G. Brewer, and D. G. Heppner. 1999. Comparison of a rapid field immunochromatographic test to expert microscopy for the detection of *Plasmodium falciparum* asexual parasitemia in Thailand. *Acta Trop.* **73**:263–273.
 40. World Health Organization. 2000. New perspectives: malaria diagnosis. Report of a Joint WHO/USAID informal consultation. World Health Organization, Geneva, Switzerland.
 41. Zhong, K. J., C. J. Salas, R. Shafer, A. Gubanov, R. A. Gasser, Jr., A. J. Magill, J. R. Forney, and K. C. Kain. 2001. Comparison of IsoCode STIX and FTA Gene Guard collection matrices as whole-blood storage and processing devices for diagnosis of malaria by PCR. *J. Clin. Microbiol.* **39**:1195–1196.
 42. Zhou, M., Q. Liu, C. Wongsrichanalai, W. Suwonkerd, K. Panart, S. Prajakwong, A. Pensiri, M. Kimura, H. Matsuoka, M. U. Ferreira, S. Isomura, and F. Kawamoto. 1998. High prevalence of *Plasmodium malariae* and *Plasmodium ovale* in malaria patients along the Thai-Myanmar border, as revealed by acridine orange staining and PCR-based diagnoses. *Trop. Med. Int. Health* **3**:304–312.
 43. Zweig, M. H. 1988. Evaluation of the clinical accuracy of laboratory tests. *Arch. Pathol. Lab. Med.* **112**:383–386.
 44. Zweig, M. H., and G. Campbell. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39**:561–577.