Methods

Study Design

Between March 5 and May 16, 2007, a prospective, community-based participatory research cohort quality control study was performed to assess the diagnostic accuracy of malaria microscopists at the Mae Tao Clinic in Mae Sot, Thailand.

Selection and Description of Participants

Study participants were clinic outpatients five years of age or older, including pregnant women, who presented to the clinic as new cases with fever and were referred to the clinic lab for malaria testing. Exclusion criteria were age less than five years, anti-malarial treatment and/or previous blood smears taken during the current clinic visit, and inability to understand the informed consent proceedings when conducted in the Burmese language.[†]

Verbal and written explanations of the study protocol were provided in the Burmese language, and each study participant or his/her guardian provided written informed consent with a signature or X (if the participant was unable to read or write). Subjects were informed of the risks and benefits of study participation and told that study participation was voluntary, that consenting participants would be giving two small blood samples, and that their personal identification information and test results would be kept confidential, The study was approved by the Chulalongkorn University Health Science Faculties Ethical Review Committee for Research Involving Human Subjects (Bangkok, Thailand).

[†] Most clinic patients, though hailing from diverse ethnic backgrounds (including Karen, Shan, Kachin, and Arakan), understand the Burmese language, which is used as a lingua franca among most Burmese people. However, some patients, especially those from more rural areas of Burma, only speak the language of their particular ethnic group.

Sample Collection

From March 5-31, 2007, patients with suspected malaria who presented Mae Tao Clinic medics were referred to the Mae Tao Clinic laboratory for testing. Mae Tao Clinic laboratory microscopists collected a 0.5 ml fingerprick blood sample for PCR analysis from each patient on a strip of filter paper (Whatman, Kent, UK) labeled with the patient's unique identification code. After air-drying, PCR samples were placed in individual plastic bags and stored in an airtight, closed container until the end of the four-week sample collection period.

Thin and thick blood smears were then obtained from each patient's fingerprick site, labeled with their identification code, stained with Giemsa, and read under a light microscope.

Sample Analysis

Based on Mae Tao Clinic diagnoses of the blood smears, patients were treated per normal clinic protocol (3 days of Artesunate + Mefloquine for uncomplicated *P. falciparum* single or mixed infection; 3 days of Chloroquine + Primaquine for uncomplicated *P. vivax, P. malariae,* or *P. ovale* single or mixed infection)[6]. All blood smears taken over the 4-week sample-collection period were separated into "malaria positive" or "malaria negative" categories after initial reading and stored in corresponding slide boxes until the end of the sample-collection period.

Following sample collection, all of the "malaria positive" slides and a randomly-selected subset of "malaria negative" slides were chosen for inclusion in the final quality control study analysis (see "Statistical Analysis" section for more details). These positive and negative slides were mixed together, placed in chronological order of when they were collected, and transported to Bangkok for microscopic analysis at the Thai Ministry of Public Health. Subsequently, PCR samples corresponding to the blood smears selected for quality control

analysis were identified, separated, and transported to Bangkok for PCR analysis at the Malaria Unit, Institute of Health Research, Chulalongkorn University.

Microscopy

From April 5-May 4, 2007, expert microscopists from the Thai Ministry of Public Health [MoPH], blinded to Mae Tao Clinic results, examined and interpreted the blood smears transported from the Mae Tao Clinic [MTC]. Any blood smears for which MoPH results were discordant with MTC results (as determined by the principal investigator), were read a second time by a different MoPH microscopist who was aware that the initial results were discordant but did not know the nature of the discordant results. The second reading was taken as the final MoPH result.

Slide Quality

In addition to interpreting MTC blood smears, the MoPH was also asked to provide comments about thick smear quality, thick smear color, and thin smear quality. In order to evaluate the quality and color of each smear, a three-tiered rating scale was developed (Table 1). When determining which comments to ascribe to a slide, MoPH microscopists evaluated certain descriptive criteria (outlined in the "Smear Quality Description" and "Smear Color Description" boxes) as part of their protocol. Based on how well these criteria were met, a comment for each of the three categories was given. These comments (outlined in the "MoPH Comments" boxes) were then grouped into a 3-tiered rating scale ("Good," "Average," and "Poor") based on their subjective significance. Each slide then received one of these three grades for each of the three parameters described.

Table 1: Rating scale for evaluating smear quality and color

	- Standard - Good - Very good	- Average - Usable	- Below standard - Needs improvement - Poor - Too thin - Too thick
	- Appropriate thickness - Covers enough area (1 cm^2) - Little to no Giemsa residue - Little to no artifact - For thick smears, 10-20 WBC per field	- Moderate Giemsa residue - Moderate artifact - Does not cover appropriate area - For thick smears, 5-10 WBC per field	- Too much artifact - Too many precipitins - Too much or too little blood - Blood film spread unevenly - Blood rubbed off - Autofixation - For thick smears, ≤ 5 WBC per field - Too much Giemsa residue/dust
	- Pinkish-blue color	- Blue color	- Too red - Too blue

DNA Extraction & Purification[†]

From April 12-May 16, 2007, DNA extraction and PCR analysis of the blood samples from Mae Tao Clinic were performed at the Malaria Unit, Institute of Health Research, Chulalongkorn University according to protocol outlined by Georges Snounou et al (1993)[10].

DNA extraction was done using Chelex® 100 [ion exchange] Resin (Bio-Rad, Hercules, California, USA). To prepare the resin for each sample, 50 µl of a 20% Chelex® 100 stock solution in water were diluted with 150 µl of deionized water, the mixture then vortexed and heated in a heating box (Barnstead/Thermolyne, Iowa, USA) at 100°C for 35 min.

A 3 mm strip from each blood-saturated filter paper sample was cut and soaked in 1 ml phosphate-buffered saline (PBS) overnight at 4°C. Parasites and unlysed erythrocytes were pelleted on the paper strip by brief centrifugation (9 sec, "Quick Run" mode). The supernatant was discarded and the paper resuspended in 0.5 mL of 0.5% saponin in PBS solution. Samples were refrigerated for 10 min at 4°C while RBC lysis occurred. Parasites and white blood cells on the paper were recovered by centrifugation as above and the supernatant discarded. Each paper strip was rinsed with 1 ml PBS and centrifuged again as above, and the supernatant discarded. The paper was resuspended in 1 ml PBS and refrigerated for 30 min at 4°C to allow saponin washout. Each paper strip was extracted and immersed in 200 µl of the Chelex® 100 solution prepared in the first step. The paper and Chelex® 100 solution were vortexed for 30 sec and heated in the heating box at 100°C for 10 min. At the 5- min halfway point of heating, each sample was vortexed again for 10 sec and returned to the heating box. After the complete 10 min heating period, each sample was vortexed for 20 sec to allow DNA extraction from the paper into the solution. The paper and excess Chelex® 100 resin were separated from the DNA-containing supernatant via centrifugation (3 min, 13,000 RPM). The supernatant from each sample was then transferred to another tube and centrifuged again as above to allow precipitation of additional excess resin. The purified DNA-containing supernatant for each sample was then transferred to a final tube and stored at -20°C to be used as the DNA template for subsequent PCR amplification.

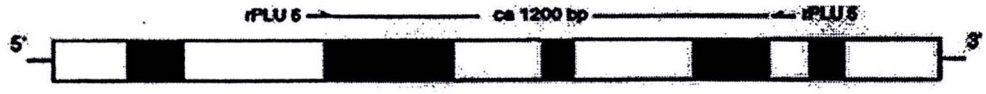
DNA templates used as positive controls were obtained from malaria parasites whose species identity had been confirmed by expert microscopy and PCR performed at the Malaria Unit, Institute of Health Research, Chulalongkorn University. The DNA was purified using the phenol-chloroform extraction protocol developed by Georges Snounou (1994).[11]

PCR Amplification

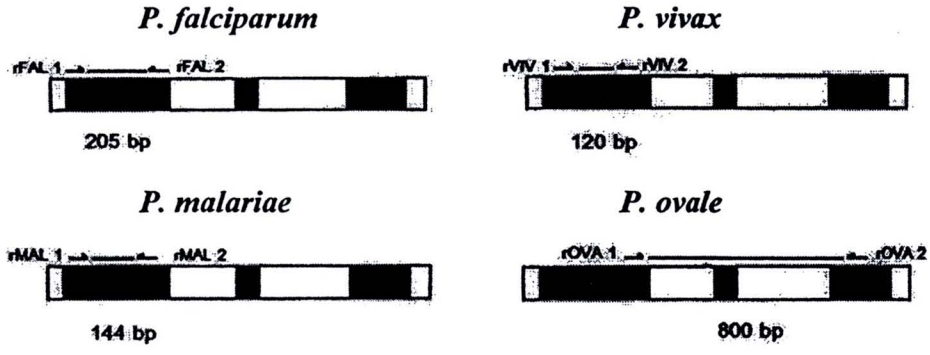
PCR amplification of Plasmodium DNA was performed using the nested PCR technique outlined by Georges Snounou et al (1993).[12] The protocol used to amplify the desired Plasmodium genes consisted of 2 reactions: 1) Amplification of the genus-specific Plasmodium gene using purified DNA from each sample as the DNA template; and 2) Amplification of each species-specific Plasmodium gene, using the DNA product from Reaction 1 as the DNA template. For the first reaction, the genus-specific primer pair rPLU5 and rPLU6 (Sigma Proligo, Singapore), was used to amplify the 1.2 kb gene coding for a Plasmodium ssrRNA sequence common to all Plasmodium species. For the second reaction, species-specific primer pairs (rFAL1/rFAL2, rVIV1/rVIV2, rMAL1/rMAL2, and rOVA1/rOVA2, Sigma Proligo, Singapore) were used to detect the presence of each Plasmodium species within the DNA product obtained in Reaction 1. This species-specific reaction was run separately for each sample. A specific ssrRNA gene product from Reaction 2 (205 bp for *P. falciparum*, 120 bp for *P. vivax*, 144 bp for *P. malariae*, and 800 bp for *P. ovale* as shown in Figure 1) was only obtained if DNA from the corresponding species was present in the reaction--i.e. a positive infection from that species in a given sample. Neither Reaction 1 nor Reaction 2 yielded any DNA product if there was no Plasmodium DNA in a given sample—i.e. a negative infection.

All PCR reactions were performed in a total volume of 20 µl. Each reaction was carried out using an AmpliBuffer (Vivantis, Chino, California, USA) containing 500 mM KCl, 100 M Tris-HCl (pH 9.1), and 0.1% Triton™ X-100; 2 mM $MgCl_2$; 125 µM of deoxyribonucleotide triphosphates; 250 nM of each oligonucleotide primer, and 0.44 units *Taq* DNA Polymerase (Vivantis, Chino, California, USA). For Reaction 1, 1 µl of purified DNA template from each sample was used to amplify the 1.2 kb genus-specific *Plasmodium* gene if present. For Reaction 2, a 1-µl aliquot of the DNA product obtained from Reaction 1 was used as the DNA template for each of the 4 separate reactions aimed to detect the species-specific *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* genes if present.

a) First amplification reaction



b) Second amplification reaction



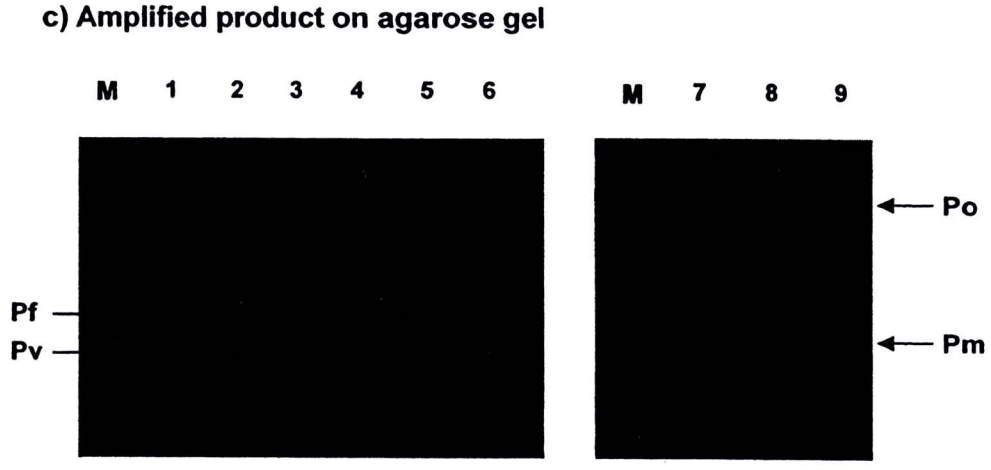


Figure 1: Schematic representation of *Plasmodium* ssrRNA genes; a) the first reaction for genus specific amplification, b) the second reaction for species specific amplification. The black boxes represent variable sequences unique to each species (Modified from Snounou et al., 1993), c) the PCR products were separated on agarose gels: M stand for 100 bp DNA ladder, lane 1 presented positive control for *P. falciparum* and *P. vivax*, lane 2-6 presented the PCR products from the samples, lane 7 presented positive control for *P. ovale* and *P.malariae*, lane 8-9 presented the PCR products from the samples.

For Reaction 1, the PCR amplification program parameters were as follows. Step 1: Initial denaturing at 95°C for 4 min; Step 2: Further denaturing at 95°C for 1 min; Step 3: Annealing at 58°C for 2 min; Step 4: Extension at 72°C, 2 min; Step 5: Final extension at 72°C, 2 min. Steps 2-4 were cycled 25 times in between steps 1 and 5. At the end of the program, the temperature was reduced to 20°C. For Reaction 2, the program parameters and cooling temperature remained the same, except steps 2-4 were cycled 30 times in between steps 1 and 5.

PCR Product Analysis

3% NuSieve agarose/Agarose (3:1) gels were made using 2.25 g NuSieve agarose (FMC BioProducts, Rockland, USA) and 0.75 g agarose (USB, Cleveland, USA) dissolved in 100 ml

1x TBE buffer (Tris, boric acid, EDTA pH 8.0), 1 µl loading dye, and 2 µl DNA product were loaded into each gel well, with a 100 kb marker and positive controls loaded into the first and second wells in each column, respectively. For most electrophoresis procedures, DNA products and positive controls from the *P. falciparum* and *P. vivax* species for each sample were loaded together in the same well, as were the DNA products and positive controls from the *P. malariae* and *P. ovale* species for each sample, in order to save time and conserve materials. The gels were immersed in 1x TBE buffer, and amplification products were electrophoresed at 70 mV. The gels were then stained in ethidium bromide and the DNA products visualized under ultraviolet light.

PCR technicians were blinded to MTC and MoPH results. Any samples for which PCR results disagreed with MTC or MoPH microscopy results (as determined by the principal investigator) underwent PCR a second time. If the second PCR result was discordant with the first PCR result, the sample underwent PCR a third time. PCR technicians were aware that previous results were discordant but did not know the nature of the discordant results. The final PCR result was considered the one that appeared 2 out of 3 times.

Statistical Analysis

Sample size was calculated based on a desired precision level of $\leq .03$ for a 95% confidence interval and expected sensitivity $\geq 95\%$, using the equation:

$$n = \frac{Z^2 p (1-p)}{L^2}$$

where n = sample size of positive *or* negative selected slides, Z = 1.96, p = estimated sensitivity (0.95), and L = desired precision level (0.03). From this equation, $202.6 \approx 203$ was

the desired sample size for positive *or* negative slides; therefore, 203 slides were randomly selected from the total number of 626 eligible negative slides (approximately 1 out of 3). Random selection was accomplished by extracting and keeping every 3rd slide from the "malaria negative" slide box without regard to any characterizing features such as perceived slide quality, date of smear collection, or patient identification information. Since the total number of eligible positive slides was only 136, all positive slides were included in study analysis in order to provide an adequate proportion of positive and negative samples.[13, 14]

Frequencies and cross-tabulations for MTC, MoPH, and PCR results were obtained using SPSS v. 13 for Windows (SPSS, Chicago, Illinois, USA). From these, the sensitivity, specificity, positive predictive value, and negative predictive value (± 95% confidence intervals) were calculated for the MTC and MoPH using standard formulas, with PCR as the reference standard, and again for the MTC using final MoPH interpretations as the reference standard. To assess what these quality indicators for the MoPH would have been had MoPH microscopists not had the opportunity to re-interpret discordant results, these four parameters were also calculated for the initial MoPH interpretation, using PCR as the reference standard. P-values were calculated to compare frequencies and quality indicators among all three groups, using an interactive chi-square calculation tool developed by Preacher (2001)[15]. Agreement of diagnostic performance among all three groups was compared using an unweighted kappa statistic. Interpretation of kappa values was based on the Altman (1991) classification.[16]

Possible results for malaria diagnosis included no infection ("negative"), a single-species infection (e.g. *P. falciparum*), or a mixed-species infection (e.g. *P. falciparum* + *P. vivax*). A microscopy result was considered "accurate" if it matched the PCR result exactly; therefore, if

a microscopist interpreted a sample as *P. falciparum* that was interpreted as *P. falciparum* + *P. vivax* by PCR, the microscopist's interpretation would not be considered accurate, even though one of the species present was correctly identified. Similarly, results were considered to be "in agreement" if they matched exactly; simply diagnosing one species correctly for a mixed infection was not sufficient to constitute agreement.

Errors made by MTC and MoPH microscopists were described, quantified, and correlated with the perceived quality of the misdiagnosed slide. P-values were calculated to compare error counts for the MTC and MoPH.

Since the above summary statistics are less important than the clinical implications for the patients to which they refer, MTC and MoPH errors resulting in incorrect therapy for *P. falciparum* and/or *P. vivax* were identified. To examine the clinical consequences of an inaccurate diagnosis, patients who did not receive adequate or appropriate treatment as a result of these errors were followed up to see whether they ever were properly diagnosed and treated at a subsequent clinic visit.