Discussions and Recommendations

Discussion

This study showed that the Mae Tao Clinic and the Thai Ministry of Public Health diagnosed malaria with comparable accuracy using microscopy. Even though MTC microscopists made almost twice as many overall errors as MoPH microscopists, most of these errors were species-specific and therefore did not result in any statistically significant difference between the overall main quality indicators for both parties. In particular, the sensitivity and specificity obtained by both groups exceeded World Health Organization Southeast Asia Regional Office [WHO/SEARO] standards for microscopy competency at the regional level (desired sensitivity $\geq$ 90%; desired specificity $\geq$ 95%).[17] Though located in a remote field area, Mae Tao Clinic performed well above the minimal expected standards for a primary diagnostic site (desired sensitivity $\geq$ 80%; desired specificity $\geq$ 85%).[17] Moreover, these results underestimate the performance of MTC and MoPH microscopists, since PCR detects malaria parasitemia at levels five times lower (1-5 parasite/mm^3 blood) than that detectable by the best microscopist[9] (i.e. yielding positive results for subpatent parasitemia). Thus, using PCR as a reference standard will underestimate MTC and MoPH sensitivity and negative predictive value. For this reason, while PCR is an appropriate reference standard for assessing diagnostic test performance, it is not usually used as the reference standard for assessing microscopy quality. Accordingly, it is likely that the MoPH, if compared against a non-PCR gold standard, would meet the WHO/SEARO standards for microscopy at the national level, as it should (desired sensitivity and specificity $\geq$ 95%).[17]

When comparing species-specific and negative/positive agreement among results, the Mae Tao Clinic had good agreement with PCR (kappa = 0.799 (95% CI 0.723-0.874)) and very good agreement with MoPH results (kappa = .827 (95% CI 0.758-0.896)). The MoPH had very good agreement with PCR (kappa = 0.889 (95% CI 0.831-0.947)). MTC microscopists tended to overdiagnose *P. falciparum*, especially in cases of single *P. vivax* infections, to a greater extent than the MoPH. Even though such errors result in unnecessary administration of the more costly artesunate-based regimen and contribute to the problem of resistance against this "last defense" drug, they are less harmful than *P. falciparum* false-negatives, given the potential clinical severity of a *P. falciparum* infection.

Mae Tao Clinic accuracy could have been affected by several factors. At the time of the study, the Mae Tao Clinic laboratory team consisted of 23 Burmese technicians whose microscopy experience ranged from a few months to over 10 years. There is a high level of turnover among staff, since many return to Burma, are resettled to another country, or cannot maintain a long-term position in the laboratory due to personal or political reasons. Consequently, it is difficult to build the expertise level of each team member and ensure a reliable, constant standard of quality. In contrast, the Ministry of Public Health laboratory team consisted of four expert technicians who have each been performing quality control microscopy for the whole of Thailand for over 10 years, following intensive mentored training, strong peer support and interaction, and adhering to a strict protocol outlined by the WHO. Despite the potentially considerable difference in skill level between the two groups, the Mae Tao Clinic performed well.

It is also noteworthy that Mae Tao Clinic staff only had one opportunity to read and diagnose each blood smear, whereas the MoPH had the chance to re-evaluate 38 slides that

were originally discordant with MTC results. Of these, 8 results were changed to a new result, 7 of which agreed with the PCR result. These changes between the initial and final MoPH interpretations are significant, as the MoPH's sensitivity would have been reduced to 91.9% (95% CI, 87.6-96.2) from 93.2% (95% CI, 89.2-97.2), and specificity would have been reduced to 98.4% (95% CI, 96.1-99.9) from 98.9% (95% CI, 97.4-100.0) had they not had the opportunity to re-examine those slides.

The high percentage of misdiagnosed slides with poor quality thick and/or thin smears may indicate that sub-par smear quality impairs the ability of microscopists to diagnose accurately, or it may simply parallel the larger overall proportion of poor quality smears present. However, it does appear that "good" thick or thin smear quality was not a guarantee of an accurate reading. The fact that a larger proportion of MoPH misdiagnosed slides had poor quality thin smears compared to MTC misdiagnosed slides may indicate that that poor thin smear quality was conducive to MoPH errors, whereas MTC errors in reading thin smears may be more attributable to skill level. Despite these ambiguities, improving smearing techniques and stain quality through regular trainings led by a reliable outside agency adhering to WHO protocol would likely improve MTC baseline accuracy in interpreting blood smears. In addition, though collected blood smears were stored appropriately, age and elemental exposure might have caused smear degradation over the four-week sample collection period so that slide quality was diminished by the time they arrived at the MoPH, resulting in a greater number of MoPH errors.

Since over half the misdiagnosed slides for both groups were "dirty" or "very dirty," efforts to reduce residue buildup on slides would probably enhance diagnostic accuracy. Suggested methods for this include washing new slides before use (to eliminate the oily film

that attracts dirt), and using filtered water to wash slides and mix Giemsa stain (to avoid mineral deposition). Though the presence of precipitins from oxidized Giemsa were present on only a few misdiagnosed slides, filtering Giemsa stain and ensuring that ensuring that each batch of Giemsa is used within an hour of mixing would minimize oxidation and prevent precipitins from potentially affecting slide diagnosis.[18]

In this study, MTC and MoPH microsopist performance may have been underestimated, as it was assumed that PCR, the "gold standard" reference test, had 100% accuracy. It is well documented that PCR yields positive results for a longer time after treatment than does microscopy, which in some cases results in false positive diagnosis of an infection that may be clinically insignificant or does not warrant treatment. Srinivasan et al (2000) demonstrated that PCR detected plasmodium DNA in 12 of 13 patients that were apparently aparasitemic by microscopy 2-5 days after treatment.[19] While these findings were thought to represent circulating *plasmodium* DNA from possibly non-viable parasites killed by appropriate drug treatment, Jarra et al (1998) asserts that only viable parasites are detectable by PCR.[20] If Jarra et al are correct, then PCR might still show a positive result for patients with subpatent parasitemia, a partially treated/recrudescent infection, live but damaged parasites that are unable to cause infection, or gametocytemia (presence of sexual forms of the parasite that are transmissible but do not cause clinically significant infection or respond to drug therapy).[21,22,23] Since the latter two options qualify as a clinically "negative" infection, a positive PCR result for these reasons would be considered a "false positive" result that microscopists would likely diagnose as "negative."

In addition, though PCR protocol was performed carefully and discordant results repeated, sample contamination could have resulted in false-negative or false-positive PCR findings,

which would have mistakenly attributed error to the microscopists. The fact that 49 PCR samples had to undergo amplification a third time because of discordance between first and second interpretations indicates that the potential for error is considerable. In addition to sample contamination and cross-contamination, possible sources of PCR error include skill level of PCR technicians (one who was a relative novice and one who had several years' experience), mechanical error from the PCR or gel electrophoresis equipment that caused reduced sequence amplification or inconclusive visualization of PCR products, and misinterpreting a PCR product on the gel as a false positive or false negative. In addition to the fact that PCR must be performed under near-sterile conditions, the many factors contributing to potential PCR errors emphasize why PCR is unrealistic under field conditions and should not be considered the gold diagnostic standard when ideal laboratory conditions cannot be upheld.

Most patients untreated for *P. falciparum* or *P. vivax* infection at their initial Mae Tao Clinic evaluation (53.8% for untreated *Pf*; 75% for untreated *Pv*) never received proper treatment and were lost to follow-up. However, almost half of patients with initially untreated *P. falciparum* and a quarter of those with initially untreated *P. vivax* returned to the clinic for follow-up and received treatment (albeit sometimes for the wrong species) that ultimately yielded a negative blood smear at a subsequent clinic visit. These observations underscore the importance of being able to access health care services for follow-up visits (which many of the transient MTC patients cannot) and indicate that some *P. falciparum* infections are still susceptible to chloroquine and primaquine, the treatment for *P. vivax*. Despite the favorable outcomes in this subset of patients, however, it is still of paramount importance to diagnose

malaria infections correctly the first time and assume that all *P. falciparum* infections are resistant to chloroquine/primaquine until proven otherwise.

Despite the increasing use of rapid diagnostic testing (RDT), an immunoglobulin-based assay to diagnose malaria antigens, in more remote clinical settings over the years, the WHO still recommends microscopy as the first-line laboratory diagnosis for malaria due to its superior sensitivity and ability to detect multiple malaria species and other conditions .[24] Even though RDTs are preferred in settings where "good microscopy cannot be maintained or is non-existent,"[25] the WHO stresses the importance of extending and strengthening good microscopy facilities "closer to the periphery of health services" in order to more accurately differentiate species and thus dictate drug-specific treatment for *P. falciparum*.[24] Nevertheless, there are those who believe that RDT's, whose results do not depend on subjective interpretation, reliable equipment maintenance, or skill level of the administrator, should be used more widely in a variety of field settings, and that a febrile patient with a negative *P. falciparum* RDT result should be treated empirically for *P. vivax*.

While debates about the best paradigm for malaria diagnosis in the periphery continue, microscopy is still the preferred diagnostic modality along the Thai-Burma border. As one such organization in this region, the Mae Tao Clinic aims to serve as training hub for other clinics across the border in Burma that would like to develop and strengthen their own microscopy facilities. Results from this study demonstrate that the Mae Tao Clinic has been able to implement an adequate malaria microscopy facility in a resource-limited field setting and is thus qualified to take on this responsibility. To further evaluate which diagnostic modalities are most advantageous in the varied clinical settings along the Thai-Burma border, a similar quality control study incorporating blinded comparison of microscopy, RDTs, and

PCR at a field site like the Mae Tao Clinic might be useful. In addition, given the ramifications of missed diagnosis and inappropriate treatment, a standardized, reliable network of monitoring, evaluation, and quality control strategies should be implemented among all health service providers in this highly endemic region.