

Perspectives

Rapid Malaria Tests: Where Do We Go After 20 Years?

Michael T. Makler and Robert C. Piper*

Flow Inc., Portland, Oregon; Physiology Department, University of Iowa, Iowa City, Iowa

Abstract. Great advances have been made in developing rapid diagnostic tests (RDTs) for diagnosing malaria. To date, RDTs present an exceedingly practical format for malaria diagnosis that outperforms traditional microscopy and more experimental next generation devices in the development pipeline. However, although use of such tests is accepted in principle, their actual use has lagged. Furthermore, study of how these tests perform, what their limitations are, and how to work with these limitations to still use them effectively has stagnated. We propose that the study and implementation of such RDTs should be aggressively advanced and propose a series of questions that can guide efforts.

INTRODUCTION

We believe that good health care should be everyone’s right. This means that persons in the developing world, just like those in richer countries, should have access to basic medical care, which is founded on the best science and practice. The threat of malaria poses a compelling example because although this widespread disease is relatively restricted to the developing world, its treatment warrants the best efforts that modern medicine can offer. One of the central tenants of modern medicine is accurate and timely diagnosis. However, after trips to Africa, Central and South America, Thailand, and Madagascar, countries where malaria is endemic and access to medical care is difficult, we came to understand that physical conditions and limited training meant that diagnosing malaria by microscopy was an uncertain endeavor. Blood samples could be compromised when delivered after hours on a bicycle in the hot sun. Microscopes broke down, parts were not always available, and often there were no technicians who could read the language in which repair manuals were written.

It takes great skill and years of experience to learn to accurately read a malaria slide. It is easy to make mistakes. Microscopes are costly, as is training. When technicians misread a slide and persons are treated for the wrong form of malaria, or treated when they do not have viable parasites, drug resistance is amplified. Thus, we decided to develop a low tech color-based rapid malaria test that could be read as easily as an over-the-counter pregnancy test. Other investigators have developed similar technology based on detecting other parasite proteins. Through commercialization, there are now several rapid malaria tests (RDTs) available. Although methods such as the polymerase chain reaction remain a future alternative to these tests, the practical, accessible, and economical nature of RDTs make them the only currently viable supplement to or replacement of microscopy-based diagnosis.

Initially, the use of RDTs met stiff resistance in the parasitology community; the first concern was cost, with a second concern for performance. A distant third concern was whether effective treatment of malaria would truly benefit from accurate diagnosis. A variety of reports from a broad spectrum of policy makers have acknowledged that RDTs may have their place because expert microscopy in malaria-endemic coun-

tries is hard to establish and maintain and the cost of RDTs has been greatly reduced.1

Are RDTs cost effective? Depending on the scale, RDTs can be purchased for $0.50 U.S. dollars, with the promise that the price might be reduced to $0.30–0.40 if worldwide demand fostered the economy of scale. Cost–benefit analyses of RDT versus presumptive therapy define a narrow window of advantage.2–5 Where Plasmodium spp. are less prevalent, the benefit increases and it becomes increasingly important to correctly identify these situations because of loss of immunity. However, even in highly malaria-ridden areas where the strict cost/ benefit analysis would justify RDT use by only a few cents, it is important to remember the obligation we have to bring the best of what medicine can offer, especially when malaria treatment and prevention schemes are heavily subsidized.6 Finally, if our mark of success in these areas is to make malaria rarer, establishing a diagnostic infrastructure early on will become increasingly valuable when it must be relied on more. Thus, we feel that arguing against using RDTs using cost–benefit analysis is neither mathematically compelling nor medically responsible.

There are also reservations about how well RDTs perform. Indeed, there are well-founded concerns that commercially available RDTs may lack the consistency, quality control, and performance capabilities that manufacturers claim, making their use ineffective or potentially dangerous.7,8 However, there is ample evidence that RDTs can meet and exceed benchmarks proposed for an effective diagnostic tool.9

Thus, the question becomes how do we use RDTs as an effective diagnostic and is this something we are prepared to do. We believe this is a critical issue because of the reluctance in using RDTs as a universal diagnostic tool may reflect the skepticism of what benefit malaria diagnosis really has. Efforts based on treatment (drugs and the promise of a future vaccine) and prevention (DDT and bed nets) are often emphasized and the implementation of diagnosis is often passed over. No antimalarial drug or prevention strategy is 100% effective; however, we advocate their aggressive and broad use. Why is this not the same for RDTs, even with their acknowledged limitations? Without recognizing exactly what benefits diagnostics can provide and then improving the cost or performance of existing RDTs, developing more high tech next generation diagnostics will not be helpful.

Cost and performance are not the most critical issues. Instead, we believe there should be a vigorous evaluation of what we need from diagnosis, how much we need it, and how best to meet those needs by using existing technology.
Answers are required for several questions, with the realization that they may vary from region to region. Should diagnosis of malaria routinely precede drug treatment? The alternative is that presumptive diagnosis based on patient history and other factors would be used. Should follow-up diagnosis be made to identify drug-resistance? Here, ensuring the efficacy of treatment of individual patients and identifying an emerging threat of drug-resistance to a patient population could be useful. Does detecting the malaria parasite accurately predict disease? Detecting the presence of malaria could serve as a good explanation for clinical symptoms. However, the risk is that the patient may have something else in addition to malaria, making it possible that RDT use could compromise clinical judgment in some instances.

These questions cannot be fully answered now but they can be addressed with dynamic and large scale empirical analysis. Because the cost and performance of RDTs has been sufficiently honed and developed, we now have to opportunity to truly investigate these issues and find out just how effective diagnosis can be in contributing to the fight against malaria.

**EVALUATION OF PRINCIPLE AND PRACTICE**

More than 150 studies on RDTs have been performed. Interpreting what these data mean requires a basic understanding of the scientific principles supporting the use of a given antigen as an indicator of malaria infection and an appreciation for how such principles are put to practice in a given test kit. These are two different things. However, they are often convoluted in the literature. The result is a hodgepodge of redundant studies that provide limited conclusions and hinder deployment of diagnostics against malaria.

Questions about the basic scientific principles include the following. Does antigen show variability among strains or species of *Plasmodium* that would prevent RDTs from detecting subsets of parasites? Does the antigen level correlate with disease? If so, how well and what stages of the parasite life cycle are relevant? Is the antigen stable in blood samples so that stored and fresh samples can be accurately evaluated? These questions are important as they describe the potential and limitations of any RDT based on a particular core technology.

Questions about practice are confined to individual RDT test kits and typically assess aspects of performance. Although they may indicate conclusions about the test principle, they can only reliably report on a particular batch of tests, and the particular conditions used for evaluation. Despite the many studies that evaluate particular RDTs, efforts are seldom made to reconcile results with previous data, and only a few studies have a rigorous plan in place to determine what the basis is for discrepant results or for quality control of the RDT test procedures. There has also been a tendency to extrapolate too many general conclusions from these studies, especially conclusions about the principle of the core technologies.

**PRINCIPLES OF HRP-2 AND pLDH**

The core characteristics of histidine-rich protein-2 (HRP-2) and *Plasmodium* lactate dehydrogenase (pLDH) as antigens for detection of malaria have been known for sometime. One of the first publications of HRP-2 tests showed that the ParaSight-F test had sensitivity of 88.9% and a specificity of 87.5%. Detectable antigenemia in a group of 40 patients declined after with sulfadoxine-pyrimethamine, and by 10 days post-treatment, all but four persons were antigen free. These four persons, although clear of peripheral parasitemias, remained antigenemic for 14 days. A study by Beadle and others documented a threshold effect of HRP-2 detection showing that parasitemias of approximately 60 parasites/µL were detected, but far less sensitivity was observed for lower parasitemias. These studies established the threshold detection for parasites by using HRP-2 at parasitemias of approximately 100 parasites/µL. They also established that HRP-2 can sometimes persist after peripheral parasitemias are cleared.

Most of HRP-2 is found inside the parasite, and levels of HRP-2 largely follow the number and developmental stage of the parasite, with the caveat of persistent antigenemia. Histidine-rich protein 2 is specific for *P. falciparum*, yet there is variability in HRP-2 among different *P. falciparum* strains, with some strains lacking HRP-2. A number of critical questions concerning the principle of using HRP-2 as an accurate biomarker for malaria remain, and progress in RDT research will rely on addressing these questions.

Despite good molecular characterization of the HRP-2 gene, it is unclear which histidine-rich repeats in HRP-2 are specifically detected by the commercial HRP-2 diagnostic kits. Deletion of HRP-2 could obviously undermine the utility of HRP-2-based diagnostics. However, it is unclear how common or geographically restricted are HRP-2 gene deletions what types of selection pressures might expand their prevalence. A critical question is how HRP-2 persists in the blood and what to do about it. Histidine-rich protein 2 is a vacuolar protein, which makes it fairly stable in harsh environments. Is non-parasite–associated HRP-2 trapped within macrophages and other immune cells or is it trapped in other blood compartments. Do high levels of non-parasite–associated HRP-2 correlate with initial parasitemias, natural immunity to malaria, tendency to cerebral malaria, or risk of recrudescence? Answers to these questions will help inform the best ways to use HRP-2-based tests most effectively despite their limitations. Although persistence of HRP-2 presents an obvious weakness when diagnosing adults in heavily malarious regions, it could be advantageous in areas where malaria is seasonal or rare and where detection of current or recent malaria infection would still garner treatment or increased clinical attention. Just how advantageous or disadvantageous antigen persistence poses remains a matter of speculation that deserves empirical evidence to determine how much of a limitation it really poses.

The characteristics of the core technology for pLDH have also been outlined early on. These studies used dipsticks and an enzyme-linked immunosorbent assay (ELISA)–type procedure to evaluate the characteristics of pLDH. They showed good correlation between the level of parasitemia and the level of pLDH enzyme activity and protein. They also showed that pLDH dipsticks had a threshold effect of approximately 200 parasites/µL. Levels of pLDH were more consistent with peripheral parasitemias than HRP-2, an effect that was most obvious during the course of drug treatment. Additional studies showed that parasites in patient blood samples had to be viable for there to be detectable pLDH.

*Plasmodium* LDH functions in the glycolytic cycle, an essential process that ensures the presence of pLDH in each parasite, unlike HRP-2. Studies on sequence variability among pLDH isoforms from different strains showed low diversity,
which suggested that the antigen is a constant.15,22 Also, pLDH is expressed at levels that correlate with parasite cell size and are present in all blood forms including gametocytes.19 This latter observation has caused other investigators to erroneously conclude that the presence of pLDH in gametocytes renders pLDH-based RDTs incapable of analyzing therapy and parasite clearance.23–26 Typically, however, the residual gametocytemia resulting after malarial treatment is low and often below the threshold for detection of viable parasites by pLDH-based dipsticks.27,32

Moreover, artesiminin-based combination therapy, the antimalarial drug of choice, results in the lowest post-treatment gametocytemias.19 These factors account for the success in following therapy in patients in the field.31–39 High post-treatment gametocytes can indicate the emergence of drug-resistant populations40,41 and recrudescence infections.42 Furthermore, gametocytemias observed in patients before treatment can be associated with higher levels of asexual parasites, and a sizable proportion of patients diagnosed by microscopy to have gametocytes only prior to treatment may have asexual parasites below detection limits of the microscopist.43 Thus, we question the wisdom of excluding high gametocytemias of viable parasites from treatment, which would not occur if clinical decisions are based primarily on pLDH detection. Moreover, without large-scale studies that base treatment decisions on the presence of pLDH, the true impact of such issues will remain speculative.

The relative levels of pLDH, HRP-2, and other antigens is also an important issue because RDTs have a theoretical limit of sensitivity based on the number of antigens present in a sample. Studies using commercially available ELISAs estimate that the number of pLDH molecules in intraerythrocytic parasites is comparable to the level of HRP-2 proteins, although differences in the sensitivity of the two assays and characterization of the enzyme standards prevented exact comparisons.44 The comparable levels of antigen indicate that the theoretical limit of pLDH versus HRP-2 detection is comparable, which is consistent with field studies that demonstrate a similar threshold of 100 parasites/μL for reliable detection. It is also important to note that levels of pLDH and HRP-2 measured by ELISA roughly correlate with parasitemia. However, this correlation is not high. The factors that contribute to variability in antigen levels will be critical to determine if we are rely on these diagnostic markers. Furthermore, they underscore the danger in extrapolating test performance from a small collection of samples.

THE PRACTICE OF RDTs

Most published studies on RDTs involve product evaluation on a small-to-medium-sized cohort of blood samples and compare one test against microscopy, which inevitably leads to discrepancies. Comparison of microscopic methods and personnel also leads to discrepancies.45–49 What these discrepancies mean in terms of the validating the utility of RDTs in general or how they comment on the effectiveness of particular antigens and formats for their detection is often not clear. Such studies continue 15 years after evaluation of the first malarial RDT.10 Although such studies are helpful in documenting the performance of a particular test in a local setting, the findings are largely anecdotal, adding little to the discussion of what such discrepancies mean to the use of RDTs in general. Additionally, merely documenting such discrepancies does not explain their cause, which would otherwise guide proper use of RDTs and identify technical problems that could be solved with further development.

Recently, a few studies provide data directly comparing various RDTs and provide methods to verify the gold standard with complementary techniques such as combining microscopy with the polymerase chain reaction. These studies are helpful because they document clear discrepancies between a given product and a thoroughly evaluated blood sample. However, the basis for such discrepancies go largely unexplored. Control procedures to ensure that the test kits arrive uncompromised and could be performed properly according to manufacturer’s instructions are seldom described even though they are a required feature of rigorous studies.48 Although speculating on the cause of discrepancies has brought to light a number of potential issues, studies rarely present a second line of experimentation to definitively identify the root causes.

A more severe problem is that such results are often over-interpreted and used to comment on the general performance of RDT principles when a much more narrow conclusion is warranted. For example, Chiodini and others49 studied the stability of a batch of pLDH tests from an unnamed manufacturer and found that they were unstable when exposed to heat. Although the authors acknowledged that their data can only be narrowly interpreted to indicate that heat stability can sometimes be a problem for tests, a number of subsequent authors have concluded that pLDH-based tests in general are inherently unstable in heat.50,51,52–56

Without delving into the cause of particular discrepancies, making general interpretations from comparative studies will be limited and confined to assessing one product over another. Although many studies of particular products conclude the that pLDH tests have greater specificity but equal or inferior sensitivity compared with HRP-2-based tests,51,53,54 studies using different products have shown both superior sensitivity and specificity compared with popular HRP-2-based tests.55 Moreover, comparison of different pLDH tests, many of which use the same pLDH antibodies, shows clear differences in product kits.50,56–58 These differences highlight the issue of quality control, which contributes to performance differences between any two products.

Efforts to systematically address quality control have been initiated at the World Health Organization (FIND program) as part of an RDT diagnostic program. This program was initiated seven years ago to systematically compare different products. In large part, this program was initiated to address the shortcomings of RDT evaluation studies that we echo here. This program is clearly needed and may supplant the type of product evaluation that currently remains the topic of published studies in academic journals. However, this program has only recently published comparative findings about available RDTs. Furthermore, it has not provided rigorous assessment of antigen content of the few blood samples it used for testing, the source and integrity of recombinant standards it uses, or validation of its methods for antigen quantification. These efforts fall seriously short of what is required to address the questions posed above with regard to the basic science behind the test principles themselves. The larger questions that relate to cause and circumstance of antigen variability, the suitability of particular antigens as malaria biomarkers, and how to prescribe adaptable methods to deploy existing tests most effectively seem well out of reach by the World Health Organization FIND program.
Work on many basic questions that can be addressed by malaria research and funding organizations will help us understand the true promise and limitation of RDTs. Cases of clear discrepancy should be taken to the next level to determine their molecular basis. Is one antigen missing or degraded? Does the ratio of HRP-2 to pLDH vary among different parasite strains or stages? Or are such discrepancies attributable to less systematic causes. Do the higher sensitivity measurements for HRP-2 in some studies really reflect the inherent ability of those tests to measure viable parasites at low concentrations, or does it reflect that low level parasitemia of people in highly malarious areas who also harbor a residual pool of HRP-2, which contributes to the lingering antigenemia that causes false-positive results for HRP-2.

Establishing what these tests are capable of and ensuring their consistent performance are important goals. However, even at this stage, the performance of such tests in practice has clearly indicated their strengths and weaknesses. Like all diagnostics, there are limitations to how well they can accurately predict disease and guide the most effective treatment. Better diagnostics are always desired and theoretically possible, but pragmatic and informed use of less refined diagnostics still has a place. For instance, LDH isoenzyme tests for heart attack have now been abandoned for more accurate Troponin I-based tests. However, no one would advocate that before development of Troponin I-based tests, physicians refrain from an LDH test simply because something better might eventually come along. One could argue that the education, infrastructure, and medical benefit of using LDH as a test for the first practical enzyme test for myocardial infarction made the evaluation, acceptance, and widespread use of Troponin I-based tests that much easier.

Similarly, adoption of RDTs with a clear understanding of their capabilities can still be effective in our general fight against malaria. Deferring the deployment of diagnostics until something radically better is developed is unfortunate in the short-term and would make it even more difficult to accept and apply new technologies in the future. The success of relatively well-researched, long-established, and commercially sustainable RDTs has been bogged down by skepticism. Given that one should worry about how well a new, untested, and relatively expensive new device fare? As important as it is to evaluate the basic science of how these tests perform, we should also better investigate how best to use what we have to be the most effective.

Several studies have begun to examine these issues with the idea that the capabilities and limitations of each diagnostic test can be used to tailor their use for the most benefit. Semi-immune adults in highly malarious may benefit from a different diagnostic test than children and pregnant women in areas of infrequent highly seasonal malaria. Models have been proposed to fit particular diagnostics to particular needs, but these models should be tested. How effective would it be to consistently and aggressively blanket a geographically definable area with RDT-guided drug use for symptomatic and asymptomatic patients? Would the cost-benefit ratio shrink when real lives are saved? Persons have speculated that treating everyone with a positive HRP-2 test result would be a massive waste of drugs, but is this really true? How many drugs would be wasted, especially if this could reduce transmission? How might pLDH be used effectively to follow drug therapy and when would patients benefit most from early detection of recrudescence? Are these tests equally effective on non-immune patients? Do these tests do a better job at detecting parasite species not ordinarily encountered by local microbiologists? Can these tests be operated by village elders, pharmacists, or end-users so that they can penetrate into areas where microscopes and treatment facilities are remote?

These are all important questions that mostly seem to be waiting until we all agree we have the perfect diagnostic. We could wait longer or figure out how to make what we have now work effectively.

Received April 20, 2009. Accepted for publication August 27, 2009.


Authors’ addresses: Michael T. Makler, Flow Inc., 6127 SW Corbett Avenue, Portland, OR 97239, E-mail: mikcatflow@aol.com. Robert C. Piper, Physiology Department, University of Iowa, 5660 Bowen Science Building, Iowa City, IA 52242, E-mail: robert-piper@uiowa.edu.

REFERENCES


