Molecular malaria diagnostics: A systematic review and meta-analysis

Johanna M. Roth, Daniël A. Korevaar, Mariska M. G. Leeflang & Pètra F. Mens

To cite this article: Johanna M. Roth, Daniël A. Korevaar, Mariska M. G. Leeflang & Pètra F. Mens (2015): Molecular malaria diagnostics: A systematic review and meta-analysis, Critical Reviews in Clinical Laboratory Sciences, DOI: 10.3109/10408363.2015.1084991

To link to this article: http://dx.doi.org/10.3109/10408363.2015.1084991

Published online: 17 Sep 2015.
Molecular malaria diagnostics: A systematic review and meta-analysis

Johanna M. Roth1, Daniel A. Korevaar2, Mariska M. G. Leeflang2, and Petra F. Mens1,3

1Royal Tropical Institute (KIT), KIT Biomedical Research, Amsterdam, The Netherlands, 2Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, The Netherlands, and 3Department of Infectious Diseases, Division of Internal Medicine, Center of Tropical Medicine and Infectious Diseases, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Accurate diagnosis of malaria is essential for identification and subsequent treatment of the disease. Currently, microscopy and rapid diagnostic tests are the most commonly used diagnostics, next to treatment based on clinical signs only. These tests are easy to deploy, but have a relatively high detection limit. With declining prevalence in many areas, there is an increasing need for more sensitive diagnostics. Molecular tools may be a suitable alternative, although costs and technical requirements currently hamper their implementation in resource-limited settings. A range of (near) point-of-care diagnostics is therefore under development, including simplifications in sample preparation, amplification and/or read-out of the test. Accuracy data, in combination with technical characteristics, are essential in determining which molecular test, if any, would be the most promising to be deployed. This review presents a comprehensive overview of the currently available molecular malaria diagnostics, ranging from well-known tests to platforms in early stages of evaluation, and systematically evaluates their published accuracy. No important difference in accuracy was found between the most commonly used PCR-based assays (conventional, nested and real-time PCR), with most of them having high sensitivity and specificity, implying that there are no reasons other than practical ones to choose one technique over the other. Loop-mediated isothermal amplification and other (novel) diagnostics appear to be highly accurate as well, with some offering potential to be used in resource-limited settings.

Keywords

Diagnostic test accuracy, loop-mediated isothermal amplification, microscopy, molecular diagnostic, Plasmodium, point-of-care, polymerase chain reaction, rapid diagnostic test

Introduction

Accurate and fast diagnosis of malaria is essential for the initiation of proper treatment. Diagnostic tests for malaria in endemic areas are now recommended as routine by the World Health Organization (WHO), in all patients suspected of malaria1. However, correct diagnosis is challenging, especially among individuals with low parasitaemia in resource-poor settings. Microscopy is still often considered to be the gold standard, but it can be time consuming and has a relatively high limit of detection (LoD), which strongly
depends on the quality of the slide and the training level of the microscopist. The LoD of Giemsa-stained thick blood film is estimated to be 4–20 parasites/μl (p/μl) when performed by expert microscopists using good quality equipment, but under field conditions 50–100 p/μl are more realistic2. Antigen-detection-based rapid diagnostic tests (RDTs) are fast and easy to perform, but often have an even higher LoD (>100 p/μl), especially if the infecting agent is non-\textit{falciparum}1,3. Furthermore, histidine-rich protein II (HRP-II)-based RDTs are unsuitable for patient follow-up after treatment, because antigens may circulate up to 28 days after the infection has been cleared, leading to false positive test results in the first month after treatment3,4.

The sensitivity of molecular diagnostics, such as polymerase chain reaction (PCR) or nucleic acid sequence based amplification (NASBA), is generally believed to be higher than that of microscopy or RDTs5. Currently, the most commonly used molecular diagnostic for malaria is PCR, with a LoD between 0.004 and 30 p/μl6. Although not yet common as routine diagnostic, PCR is increasingly being used for confirmatory diagnosis of returning travelers or migrants suspected of malaria in Western laboratories and in some (reference) laboratories in endemic areas. However, implementation in resource-poor endemic settings remains a challenge, as molecular diagnostics normally require well-equipped laboratories, expensive supplies and trained technicians.

In an attempt to make molecular tools more suitable as (near) point-of-care (PoC) tests for the diagnosis of malaria in resource-poor settings, a range of easy to perform assays are under development that circumvent some of the drawbacks of conventional molecular methods. For example, isothermal amplification technologies such as loop-mediated isothermal amplification (LAMP) and NASBA do not require expensive PCR-machines7,8. PCR-nucleic acid lateral flow immuno-assy (PCR-NALFIA) uses simple and cheap lateral flow read-out systems9. These characteristics make them potentially powerful tools, not only for case-management, but also for follow-up of patients after treatment and to detect asymptomatic carriers during screening programs.

The range of molecular tests developed so far shows high variability in performance characteristics and may differ in type and volume of input material (DNA, RNA or whole blood), target gene, (number of) species detected, primer/probe composition and concentration, amplification technique (PCR or isothermal), read-out (gel-electrophoresis, fluorescence detection, lateral flow) and whether it is qualitative or quantitative10. Differences in accuracy between assays can therefore be expected. It is important to get a clear overview of the available diagnostics, their accuracy and characteristics, to enable decisions being made on whether a test is suitable for a particular purpose (clinical practice, asymptomatic screening and/or treatment follow-up) and setting (varying from well-equipped reference laboratories to rural health facilities).

Determining and comparing accuracy of the different molecular assays is challenging, because a true gold standard does not exist so far. When evaluating PCR with microscopy as a reference standard, some false positive test results may actually be sub-microscopic low-density infections11, leading to underestimations of specificity. It can be anticipated, however, that high-quality expert microscopy will detect more low-density infections and will therefore be a better reference standard than routine field microscopy. In addition to microscopy, commonly used PCR-based assays such as nested and real-time PCR are sometimes used as reference standards to determine the accuracy of a (novel) molecular tool, as these assays are thought to have superior sensitivity and specificity. Within study comparisons with both microscopy and PCR as reference standards are valuable in estimating the number of low-density infections missed by microscopy. Occasionally tests are evaluated against a composite reference standard of molecular and/or non-molecular techniques in an attempt to obtain optimal sensitivity and specificity.

Several descriptive reviews on the molecular diagnosis of malaria have been published. A thorough overview of both traditional PCR-based diagnostics and newer techniques is given by Vasoo and Pritt10. Different PCR-based methods for the diagnosis of imported malaria are discussed by Berry et al.9 while Cordray et al. review the available molecular PoC tests for endemic settings12. However, the accuracy of currently available molecular diagnostics has never been systematically reviewed, while accuracy data are essential in determining which molecular test, if any, would be the most promising to be deployed. The aim of the present review is to provide an update on the molecular malaria diagnostics that have been developed and to systematically evaluate their published accuracy.

Methods

Eligible studies

Eligible were primary studies that assessed the accuracy of a molecular diagnostic test for the detection of malaria in clinically suspected patients recruited in malaria endemic settings, using either microscopy or a PCR-based assay as a reference test. Studies using a composite reference standard including at least microscopy or PCR were also eligible. There were no restrictions regarding year of publication. Molecular tests in any format detecting any \textit{Plasmodium} species by DNA and/or RNA amplification were included. Studies performing the molecular test under evaluation in a non-endemic setting were only included if patients had been recruited in an endemic setting. Studies that included pregnant women only were excluded, as during pregnancy malaria parasites can sequester in the placenta, complicating the detection in peripheral blood of pregnant women13. A recent systematic review describes the accuracy of molecular tools for the detection of malaria in pregnancy14. Case-control studies were also excluded, because they tend to overestimate the accuracy of the test under evaluation15.

Reference test: microscopy

Microscopy is the most accurate non-molecular test available and can detect low parasite densities when performed by a trained microscopist. The procedure for routine microscopy is described by the WHO and includes reading at least 100 high power fields by a qualified microscopist16. If an included paper made a distinction between research/expert microscopy
and routine/conventional microscopy, the first was taken as the reference standard to ensure the use of the most optimal microscopy quality available. In this review, the reported quality of microscopy was assessed and considered good when at least two independent microscopists read 100 high power fields or more. These criteria are highly similar to those in a systematic review describing the accuracy of malaria RDTs compared to microscopy\textsuperscript{17}. If these criteria were not met, or when insufficient data were reported, the quality of microscopy was considered unknown.

**Reference test: PCR**

PCR-based assays are increasingly being used as a reference standard for the detection of malaria, because of their superior sensitivity in detecting low density and mixed infections\textsuperscript{18}. Studies evaluating a molecular diagnostic against a PCR-based assay were therefore included. Comparisons between different PCRs, like nested and real-time PCR, were also eligible. Comparisons using a non-PCR-based molecular diagnostic as a reference standard were excluded, as these tests are often still in the evaluation phase and currently rarely used for reference purposes.

**Search strategy**

Relevant studies were identified by searching Medline (PubMed), EMBASE, Web of Science, the Cochrane Central Register of Controlled Trials (CENTRAL), African Index Medicus and African Journals Online (AJOL). Search terms are provided in Additional file 1. Searches were performed in January 2014 and last updated in March 2014.

**Study selection**

Study references and abstracts were exported to Reference Manager (version 12, Thomson Reuters, New York, NY). Duplicates were removed. A primary selection, based on title and abstract, was performed independently by two authors (J.R. and D.K.). Studies considered relevant by at least one of the two authors were selected and their full texts were retrieved. If the full text could not be retrieved online, authors were approached whenever contact details were available. Full papers were independently assessed for inclusion by the same two authors. Disagreements were resolved by discussion or by consulting a third author (P.M.). Authors were contacted if studies were eligible, but additional data were required to derive 2 × 2 tables. If 2 × 2 tables could not be obtained, studies were excluded.

**Data extraction**

One author (J.R.) extracted data from included studies in Review Manager (version 5.2, the Cochrane Collaboration, London, UK) and Excel (version 14.0, Microsoft, Redmond, WA), which was cross-checked by a second author (D.K.). Information on study design, study participants, reference standard, index test and data to derive 2 × 2 tables was collected.

Two authors (J.R., D.K.) independently assessed methodological quality using the QUADAS-2 tool\textsuperscript{19}, which assesses risk of bias and applicability concerns in four domains (patient selection, index test, reference standard, and flow & timing). Studies were considered to have a high risk of partial verification bias if >10% of the patients did not receive a reference standard. A high risk of differential verification bias was defined as >10% of the patients receiving a different reference standard. Risk of bias was also considered to be high when >10% of the eligible patients were not included in the analysis. Disagreements were resolved by discussion or consultation of a third author (P.M.).

**Statistical analyses**

In Review Manager, 2 × 2 tables, their estimates of sensitivity and specificity and corresponding 95% confidence intervals (CIs) were plotted in forest plots and receiver operator characteristic (ROC) space. The xtmelogit command in Stata (version 12.1, Stata, College Station, TX) was used to perform meta-analyses, with a minimum requirement of four 2 × 2 tables per index test-reference standard combination. Studies were excluded from the meta-analysis if they had a high risk of bias in one or more QUADAS-2 Risk of Bias domains. Random effect models were used by default, as diagnostic accuracy studies are expected to show a considerable amount of heterogeneity, taking into account between study as well as chance variation. To compare accuracy between tests and to further investigate sources of heterogeneity, covariates were added to the xtmelogit command. A likelihood-ratio test was used to compare the accuracies of each index test included in the meta-analysis. The quality of microscopy and geographic area (continent) were investigated as sources of heterogeneity. For other factors that were anticipated to affect accuracy, like the species detected, the protocol used and the test setting, insufficient data were reported to explore whether they were sources of heterogeneity.

**Results**

**Search results and included studies**

The search identified 13 144 articles. After removing duplicates, 7 168 titles and abstracts were screened, of which 609 studies were selected to potentially fulfill the eligibility criteria. Of these, 533 were excluded based on the full text, for the following reasons (Figure 1): asymptomatic screening (n = 179), case–control design (n = 92), no formal evaluation against eligible reference standard (n = 86), study on pregnant women or congenital malaria (n = 52), duplicates (n = 26), laboratory study on *in vitro* cultured samples (n = 23), not enough data to derive 2 × 2 tables (n = 19), non-endemic cases (n = 18), full text could not be derived (n = 16), language barrier (Chinese, Russian, Greek, Japanese) (n = 13), review, editorial, letter, case report or erratum (n = 9).

Additional file 2 lists the characteristics of the 76 studies that were finally included in the review\textsuperscript{9,20–94}. All studies were published between 1992 and 2014 and covered at least 72 different endemic locations in 32 countries. Studies were conducted in Asia (n = 44), Sub-Saharan Africa (n = 21), both Asia and Africa (n = 2) and South-America (n = 9).

The molecular tests in most studies were performed in the country where study participants were recruited (n = 56). In 12 studies, samples were sent to non-endemic countries to
perform the molecular assays. In four studies, the molecular
test was performed both in the study country and a non-
endemic setting. In two studies, the molecular test was
performed in another endemic country and in two studies the
test location was unknown.

There were 13 evaluations on *Plasmodium falciparum*
only, 4 on *Plasmodium vivax* only and 12 were genus-specific.
The remainder of the assays was multiplex systems, of which
16 detected *P. falciparum*, *P. vivax*, *Plasmodium ovale* and
*Plasmodium malariae*. None of the studies specifically
investigated the presence of *Plasmodium knowlesi* infections.
There were not enough studies to perform subgroup analyses
of the different species to determine whether a difference in
test accuracy exists between the different species, most
notably *P. falciparum* and *P. vivax*. The most referenced PCR
was that of Snounou et al.95

**Methodological quality of included studies**

The quality assessment results are presented in Figure 2. An
adequate assessment of study quality was often impossible
due to incomplete reporting.

In the domain of patient selection, most studies (n = 55)
did not state whether they enrolled a consecutive or random
sample of patients. None of the studies had a case–control
design, as this was an exclusion criterion. In four studies, both
suspected and confirmed cases were included24,26,28,71. There
were concerns about applicability in two studies that included
suspected cases instead of individual patients51,73, implying
that in the case of multiple visits during the study period, > 1
test result per patient could be included in the analysis. Only
nineteen29,34,41,45,49,52–56,68,74,75,77–79,81,82,87 studies reported
that the readers of the index test were blinded and twenty-
four9,28,29,31,41,45,46,49,50,52,54–56,64,68,70,74,77,80–82,87,88,92 stu-
dies reported that the readers of the reference test were
blinded. For the remaining studies, it was unclear whether
the index test or reference standard was performed
blinded to other test results. Only 16 studies reported
that microscopy was performed by at least two independ-
ent microscopists reading 100 fields or
more20,22,30,33,34,41,43,44,55,57,80,82,83,85,86,92. The quality of
microscopy was unknown for the other studies because the
number of examined fields was not reported, there was only
one microscopist or the number of microscopists was unclear,
or the independency of microscopists was unknown. The majority of studies \((n = 45)\) reported the use of expert, research, routine, conventional or field microscopy. There was little consensus between studies on the nomenclature and characteristics of different types of microscopy, implying that no quality-based distinction between studies can be made based on the reported type of microscopy alone. A cycle threshold discriminating between positive and negative test results is applicable only to the real-time PCR assays, which was clearly pre-specified in 5 of 17 studies evaluating a real-time PCR\(^2\),\(^3\),\(^7\),\(^4\).\(^7\),\(^8\),\(^7\),\(^4\).

In the domain of flow and timing, one study showed high risk of differential verification bias \(^5\), and three studies excluded \(>10\%\) of patients from the analysis, without providing reasons\(^3\),\(^4\),\(^6\),\(^9\). Partial verification bias was always avoided.

**Findings**

Of the 76 studies in the review, 54 were included in the meta-analysis. Summary estimates of sensitivity and specificity of these 54 studies are presented in Figures 3–7. Fourteen studies reported accuracy data of a test for which an insufficient number of \(2 \times 2\) tables was available to perform meta-analysis. These data are presented in Figures 8–10. Another eight studies were excluded from the meta-analysis because only part of the \(2 \times 2\) tables could be completed \((n = 1)\)\(^6\), or because they had a high risk of bias \((n = 7)\)\(^5\),\(^8\),\(^2\),\(^4\),\(^8\),\(^7\),\(^1\),\(^3\),\(^4\),\(^6\),\(^9\). The effect on the summary estimates of sensitivity and specificity was systematically examined and in most cases they remained essentially unchanged. In case summary estimates changed after exclusion of these studies, the differences are discussed in the sections below.

**PCR versus reference standard microscopy**

Summary estimates of sensitivity and specificity of PCR after meta-analysis and their interpretation are presented in Table 1.

**Conventional PCR**

Conventional PCR was defined as a PCR containing one or more primer sets for the detection of *Plasmodium*, based on a single reaction and a gel-electrophoresis based read-out. The assays had either a single- or a multiplex format and most often used the 18S rRNA gene, but some selected merozoite surface protein-1 (msp-1)\(^5\) or msp-2\(^3\) as a target. For five studies, the target gene was not stated\(^3\),\(^7\),\(^3\),\(^7\),\(^6\),\(^8\).
Nineteen studies evaluated conventional PCR against microscopy, but five were excluded from the meta-analysis because patients were not enrolled in a consecutive or random way \((n = 1)\)\(^{28}\), there was a high risk of differential verification bias \((n = 1)\)\(^{58}\), not all patients were included in the analysis \((n = 2)\)\(^{31,44}\), or not enough data to complete 2 \(\times\) 2 tables was available \((n = 1)\)\(^{46}\).

The summary estimates for sensitivity and specificity of the 14 studies included in the meta-analysis were 98% (95% CI: 90–99) and 94% (95% CI: 83–98), respectively (Figure 3 and Table 1). When the five studies excluded based on the quality assessment were included in the analysis, summary estimates for sensitivity and specificity were 96% (95% CI: 88–99) and 97% (95% CI: 90–99), respectively. Of the 14 studies included in the meta-analysis, one evaluated the accuracy of microscopy and used PCR as the reference standard, but presented data in such a way that sensitivity and specificity of PCR could be derived\(^{90}\). There was no difference in accuracy between studies performed in Asia\(^{22,25,30,36,38,58,73,90}\) and Africa\(^{42,53,62,76,80}\). Insufficient data were available to perform subgroup analyses on the quality of microscopy.
Real-time PCR

Real-time PCR was defined as a single- or multiplex PCR whereby detection takes place either through a non-specific fluorescent dye intercalating with double-stranded DNA, or through sequence-specific fluorescent labeled DNA probes hybridizing with their complementary sequence. Of the 17 studies using real-time PCR, eight used SYBR Green or EvaGreen dye21,30,34,40,60,74,75,87 and nine used TaqMan, minor groove binding (MGB) or fluorescence resonance energy transfer (FRET) probes32,33,37,46,47,81,86,88 for detection. The 18S rRNA gene was the most frequently used target, only one study used cytochrome-b34 and one study used four different target genes for the four species the assay detected; aquaglyceroporin (AQP), enoyl-acyl carrier protein reductase (ECCR), p25 ookinate surface protein (Po25) and circumsporozoite (CS)87. No difference in accuracy was observed when comparing dye- or probe-based assays or target genes.

One study evaluating real-time PCR against microscopy was excluded from the meta-analysis because not enough data was available to derive 2 x 2 tables46. The summary estimates for sensitivity and specificity of the remaining 16 studies included in the meta-analysis were 100% (95% CI: 98–100) and 93% (95% CI: 98–98), respectively (Figure 4 and Table 1). Of those 16 studies, 9 did not evaluate the accuracy of the real-time PCR in the original paper, but rather used it as a reference standard. However, sufficient data were presented to derive diagnostic accuracy of the real-time PCR21,32–34,40,47,60,74,75. There was no difference in accuracy between studies performed in Asia21,30,32–34,37,40,43,47,81,87 and Africa60,74,75,88, nor between studies with good and unknown quality of microscopy.

Nested PCR

Nested PCR was defined as an initial PCR targeting the genus Plasmodium, followed by separate species-specific reactions using the amplified product as input material and a gel-electrophoresis based read-out. Of the 26 studies evaluating a nested PCR against microscopy, 17 used the 18S rRNA based protocol as described by Snounou et al95,96. The other studies used either 18S rRNA24,43,48,49,78,92 or 28S rRNA47,
dihydrofolate reductase (dhfr)\textsuperscript{24} or \textit{Plasmodium falciparum} multidrug resistance protein-1 (pfmdr-1)\textsuperscript{89} as target genes. No difference in accuracy was observed between assays using different target genes.

Four studies evaluating nested PCR to microscopy were excluded because patients were not enrolled in a consecutive or random way (\(n = 1\))\textsuperscript{24}, not all patients were included in the analysis (\(n = 1\))\textsuperscript{69}, not enough data to complete 2 \(\times\) 2 tables was available (\(n = 1\))\textsuperscript{46} or the same data was presented in another study within the meta-analysis (\(n = 1\))\textsuperscript{20}.

The summary estimates for sensitivity and specificity of the 22 studies included in the meta-analysis were 99% (95% CI: 98–100) and 88% (95% CI: 80–93), respectively (Figure 5 and Table 1). When the four studies excluded based on the quality assessment were included in the analysis, summary estimates for sensitivity and specificity remained the same. Of the 22 studies included in the meta-analysis, 6 did not evaluate the accuracy of the nested PCR in the original paper, but rather used it as a reference standard and presented data to derive sensitivity and specificity\textsuperscript{23,48,49,59,64,92}. There was no difference in accuracy between studies performed in Asia\textsuperscript{30,39,43,48,49,67,78,83,92–94}, Africa\textsuperscript{52–54,59,61,64,89} and South-America\textsuperscript{23,35,72,86}. Insufficient data were available to perform subgroup analyses on the quality of microscopy.

### Direct-on-blood PCR

Four studies evaluated a direct-on-blood PCR assay with microscopy as a reference standard\textsuperscript{27,41,51,84}. These assays circumvent the need for DNA extraction and are based on either conventional or nested PCR, as defined above.

The summary estimates for sensitivity and specificity were 93% (95% CI: 72–98) and 90% (95% CI: 76–96), respectively (Figure 6 and Table 1). One study compared direct-on-blood PCR to nested PCR, next to microscopy\textsuperscript{41}. In this study, sensitivity when using microscopy as a reference standard
was 100% (95% CI: 96–100) and specificity 88% (95% CI: 76–96). When nested PCR was used as a reference standard, sensitivity remained the same and specificity increased to 94% (95% CI: 83–99).

Comparisons

The summary ROC curve in Figure 11 presents the summary estimates of conventional, real-time, nested and direct-on-blood PCR compared to microscopy. The 95% CIs overlap considerably, which may imply that the true accuracy of the four PCR-based methods is very much alike. However, conventional PCR showed considerable variation in its sensitivity across studies, ranging from 51 to 100% (Figure 3). Nested and real-time PCR on the other hand showed consistently high sensitivity, with ranges of 56% (as a single outlier) to 100% and 95 to 100%, respectively (Figures 4 and 5).

The summary sensitivity of real-time PCR is slightly higher than that of conventional PCR ($p = 0.02$). Nevertheless, the summary estimates for sensitivity are very similar, with 100%, 98% and 99% for real-time, conventional and nested PCR, respectively. The specificity of nested PCR (88%) appeared to be slightly lower than that of the other two techniques (94% and 93% for conventional and real-time PCR, respectively). Based on the available data, however, it is not possible to conclude whether this was indeed due to a larger number of false positives, or to sub-microscopic infections detected by nested PCR. The sensitivity of direct-on-blood PCR appeared to be slightly lower than that of the other three techniques, but specificity was found to be similar. There was only one study directly comparing nested PCR and
real-time PCR to microscopy and to each other. The two PCRs were in complete agreement and both had a sensitivity of 99% (95% CI: 97–100) and a specificity of 97% (95% CI: 91–100), compared to microscopy.

Four studies compared PCR to a composite reference standard, combining the results of different assays into one ‘true’ outcome. Summary estimates of sensitivity and specificity are presented in Figure 8. Composites consist of either molecular data of different PCRs or a combination of microscopy, RDT and/or PCR results. Because of their different reference standards, these studies could neither be included in the meta-analysis, nor can their accuracy be compared to each other as all four used a different composite.

Finally, no clear difference in accuracy regarding the year of publication was observed for any of the PCR-based assays. This suggests that older techniques do not necessarily perform worse than newer ones and that accuracy is not inevitably lower in older compared to more recent studies.

**LAMP versus reference standards microscopy and PCR**

LAMP is a rather novel semi-isothermal amplification technique that requires two temperatures for amplification. It has been first described by Notomi et al. Four primers, recognizing six different target sequences are used and amplification usually takes less than an hour. The read-out is based on fluorescence, turbidity or sometimes gel electrophoresis. Only a UV light source is required for a visual fluorescence based read-out, which together with the fact that no PCR machine is required, is one of the key characteristics for LAMP to be considered field applicable. However, there is still need for DNA extraction.

Seven studies evaluated LAMP against a reference standard of microscopy and five against PCR, of which three compared LAMP to both methods. One study was excluded from the meta-analysis, because patients were not enrolled in a consecutive or random way. The summary estimates for sensitivity and specificity of the six studies evaluating LAMP compared to microscopy were 98% (95% CI: 94–99) and 97% (95% CI: 91–100), respectively (Figure 7 and Table 1).

When using PCR as reference standard, of four eligible studies, two evaluated different extraction- and read-out methods on the same sample set. We chose to include the comparisons that shared most characteristics between the four studies in the meta-analysis. The comparisons using a simple heat-treatment based extraction method and a visual readout, either based on fluorescence or turbidity, were therefore included. Evaluations using commercial extraction kits and/or a read-out by gel-electrophoresis were excluded from the meta-analysis when a test based on heat-treatment extraction and/or a visual read-out was evaluated in the same study (Figure 7). Summary estimates for sensitivity and specificity of the four comparisons included in the meta-analysis were 96% (95% CI: 91–100) and 91% (95% CI: 85–99), respectively.

The summary ROC curve in Figure 12 presents the summary estimates of LAMP versus microscopy and PCR. When choosing PCR as a reference standard, the confidence interval is extremely broad, overlapping entirely with the
Table 1. What is the diagnostic accuracy of PCR (conventional, nested, real-time and direct-on-blood) and LAMP to identify *Plasmodium*?

<table>
<thead>
<tr>
<th>Index test</th>
<th>Reference standard</th>
<th>No. of participants (studies)</th>
<th>Prevalence* median (range)</th>
<th>Implications</th>
<th>Quality and comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional PCR</td>
<td>Microscopy</td>
<td>2346 (14)</td>
<td>0.56 (0.07, 0.70)</td>
<td>With a prevalence of 50%, 50 of 100 patients will have malaria, as identified by microscopy. Of these, 1 will be missed by conventional PCR (2% of 50). Of the 50 microscopy negative patients, 3 will be treated additionally based on PCR results.</td>
<td>Representative patient spectrum; blinding and quality of microscopy poorly described.</td>
</tr>
<tr>
<td>Nested PCR</td>
<td>Microscopy</td>
<td>5603 (22)</td>
<td>0.48 (0.10, 0.93)</td>
<td>With a prevalence of 50%, 50 of 100 patients will have malaria, as identified by microscopy. Of these, 1 will be missed by nested PCR. Of 50 microscopy negative patients, 7 will be treated additionally based on PCR results.</td>
<td>Representative patient spectrum; blinding poorly described; quality of microscopy poorly described.</td>
</tr>
<tr>
<td>Real-time PCR</td>
<td>Microscopy</td>
<td>4974 (16)</td>
<td>0.48 (0.06, 0.98)</td>
<td>With a prevalence of 50%, 50 of 100 patients will have malaria, as identified by microscopy. Of these, none will be missed by real-time PCR. Of 50 microscopy negative patients, 4 will be treated additionally based on PCR results.</td>
<td>Representative patient spectrum; blinding poorly described; quality of microscopy poorly described; cycle threshold poorly described.</td>
</tr>
<tr>
<td>Direct-on-blood PCR</td>
<td>Microscopy</td>
<td>1404 (4)</td>
<td>0.25 (0.16, 0.64)</td>
<td>With a prevalence of 25%, 25 of 100 patients will have malaria, as identified by microscopy. Of these, 2 will be missed by direct-on-blood PCR. Of 75 microscopy negative patients, 8 will be treated additionally based on PCR results.</td>
<td>Representative patient spectrum; blinding poorly described; quality of microscopy poorly described. No studies showed high risk of bias.</td>
</tr>
<tr>
<td>LAMP</td>
<td>Microscopy</td>
<td>1449 (6)</td>
<td>0.60 (0.30, 0.92)</td>
<td>With a prevalence of 60%, 60 of 100 patients will have malaria, as identified by microscopy. Of these, 1 will be missed by LAMP. Of 40 microscopy negative patients, 1 will be treated additionally based on LAMP results.</td>
<td>Representative patient spectrum; quality of microscopy poorly described.</td>
</tr>
<tr>
<td>LAMP</td>
<td>PCR</td>
<td>604 (4)</td>
<td>0.61 (0.50, 0.73)</td>
<td>With a prevalence of 60%, 60 of 100 patients will have malaria, as identified by PCR. Of these, 2 will be missed by LAMP. Of 40 PCR negative patients, 4 will be treated additionally based on LAMP results.</td>
<td>Representative patient spectrum.</td>
</tr>
</tbody>
</table>

The results in this table should not be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These studies are referred to in the main body of the text of the review.

*Prevalence estimates are based on the median prevalence of the comparisons included in each meta-analysis.
confidence interval of reference standard microscopy, indicating a vast amount of heterogeneity.

**Other molecular diagnostics**

Apart from the most commonly used PCR assays and LAMP, other molecular tests have been developed and evaluated, but meta-analysis could not be performed because of the limited number of published papers on these techniques. Extracted diagnostic accuracy data are presented in Figures 9 and 10. Eight different assays were evaluated against microscopy, of which three studies also compared the test under evaluation with a PCR-based assay 33,54,82.

The molecular tools in this section were developed to simplify existing assays while not compromising on accuracy, in an attempt to make them more suitable as (near) PoC diagnostics or for screening asymptomatic individuals in malaria control or elimination programs.

The PCR-NALFIA uses a lateral flow device as a read-out system instead of gel-electrophoresis, thereby reducing equipment requirements and hazardous waste. PCR-NALFIA is available as a DNA- or direct-on-blood assay, whereby the latter circumvents the need of DNA extraction 9,55,56. Sensitivity ranges from 87% (95% CI: 79–93) to 95% (95% CI: 89–98), specificity from 82% (95% CI: 76–88) to 95% (95% CI: 93–97).

A semi-nested multiplex PCR is a simplification of the well-performing nested PCR, transforming it to a single-tube assay detecting the main four *Plasmodium* species and thereby reducing time and reagents. Sensitivity ranges from 94% (95% CI: 80–99) to 98% (95% CI: 93–100), specificity from 98% (95% CI: 96–100) to 100% (95% CI: 92–100)26,29.

The RNA-hybridization assay is an isothermal direct-on-blood assay in a 96-well format, with a chemiluminescence based read-out. It requires an overnight incubation step. Reported sensitivity and specificity, compared to microscopy, are 100% (95% CI: 95–100) and 98% (95% CI: 94–100), respectively. When using PCR as a reference standard, sensitivity remains 100% (95% CI: 82–100) and specificity increases to 100% (95% CI: 97–100)33. The possible application of the RNA-hybridization assay for screening asymptomatic populations in elimination settings is specifically discussed by the authors.

The PCR-enzyme-linked immunosorbent assay (PCR-ELISA) is a genus specific PCR-based assay, whereby the target is immobilized by species-specific probes and semi-quantitatively detected by a colorimetric assay. Assay time is relatively short and no hazardous waste is produced. Reported sensitivity and specificity are 97% (95% CI: 92–99) and 96% (95% CI: 86–99), respectively 50.

Nucleic acid sequence-based amplification (NASBA) is a semi-quantitative isothermal amplification technique, with a
fluorescence-based read-out. Compared to microscopy, reported sensitivity and specificity are 100% (95% CI: 94–100) and 90% (95% CI: 86–93), respectively. When using PCR as a reference standard, reported sensitivity is 94% (95% CI: 87–98) and specificity 99% (95% CI: 96–100)\textsuperscript{54}.

Ligase detection reaction-fluorescent microsphere assay (LDR-FMA) uses genus-specific PCR amplicons as input material for a species-specific multiplex ligase detection reaction with a fluorescence-based read-out. For a detailed description of the method, see McNamara et al.\textsuperscript{97} Reported sensitivity and specificity are 91% (95% CI: 83–96) and 71% (95% CI: 65–76), respectively\textsuperscript{53}. The possible application of LDR-FMA for monitoring prevalence in low transmission areas is specifically stated.

Photo-induced electron transfer-PCR (PET-PCR) is a real-time PCR using self-quenching primers, which reduces the number of reagents as compared to other real-time PCR assays. Reported sensitivity and specificity compared to microscopy are 100% (95% CI: 72–100) and 95% (95% CI: 91–97), respectively. When using nested PCR as a reference standard, sensitivity remains 100% (95% CI: 87–100) and specificity increases to 100% (95% CI: 97–100)\textsuperscript{82}. The authors specifically state the possible use of PET-PCR for large-scale screening in control or elimination programs.

The reverse transcriptase PCR detects RNA next to DNA, to increase sensitivity. Reported sensitivity and specificity are 100% (95% CI: 94–100) and 6% (95% CI: 3–12), respectively\textsuperscript{88}. A possible application in detecting the asymptomatic reservoir within a population is mentioned.

Discussion

Key findings

This review shows that conventional, nested and real-time PCR all have a very high sensitivity for detecting malaria when using microscopy as the reference standard (98%, 99% and 100%, respectively). When comparing the summary sensitivity of conventional PCR to that of real-time PCR, a significant difference was found ($p = 0.02$). This can be explained by the fact that the 95% CIs around the summary estimates only just overlap, while real-time PCR has a very narrow CI (98–100) compared to conventional PCR (90–99). This difference is in most cases clinically negligible as both sensitivities are very high and lie close together. Specificity was lowest for nested PCR (88%) and slightly higher in conventional and real-time PCR (94% and 93%, respectively). Low-specificity values may partly be explained by the fact that microscopy is considered to be an imperfect reference standard, with a LoD higher than that of the molecular index test. Some false positive test results may therefore actually be true positives with sub-microscopic parasitaemia.

The sensitivity of the direct-on-blood PCR (93%) was found to be slightly lower than that of the other PCRs, and specificity was similar (90%). The input material could be a
possible explanation for the lower sensitivity, as blood is known to contain PCR inhibitors and DNA is likely to be more concentrated in an extracted sample than in whole blood. Direct-on-blood PCR has the valuable advantage that no DNA extraction is required, saving both time and reagents and making the step to molecular PoC diagnostics easier.

The sensitivity of LAMP versus microscopy (98%) was comparable to that of LAMP versus PCR (96%). Specificity was high when comparing LAMP to microscopy (99%), but appeared to be somewhat lower when compared to PCR (91%). LAMP was expected to show higher specificity using PCR as a reference standard, as most sub-microscopic infections will not be marked as false positive by PCR. However, it should be taken into account that the meta-analysis of LAMP versus PCR was based on only four, extremely heterogeneous studies and that this analysis consisted of comparisons using heat-treatment extraction and visual read-out only. One study showed a high agreement between the accuracy of PCR and LAMP when compared to microscopy, with LAMP performed under field conditions.

Other molecular tests that were not included in the meta-analysis were: PCR-NALFIA, semi-nested multiplex PCR, RNA hybridization assay, PCR-ELISA, NASBA, LDR-FMA, PET-PCR and reverse transcriptase PCR. Sensitivity was generally high, with a range of 87–100% when microscopy was used as a reference standard and 94–100% with reference standard PCR. Specificity was high with PCR as a reference standard (range 94–100%), but more variable for reference standard microscopy, with an outlier of 6% for reverse transcriptase PCR. This outlier could possibly be explained by the detection of RNA next to DNA. NASBA, the RNA-hybridization assay and PET-PCR were evaluated against both microscopy and real-time or nested PCR. In all three assays specificity was higher when using PCR as a reference standard, compared to reference standard microscopy (Figures 9 and 10). This confirms that the detection of sub-microscopic infections by molecular assays may explain the larger proportion of false positive test results when using microscopy as a reference standard.

Strengths and weaknesses

Taking into account that microscopy is an imperfect gold standard, it is assumed that reference laboratory microscopy is more suitable as a reference standard than routine microscopy in field settings, due to its higher sensitivity. However, there is currently no consensus in the reporting of microscopy quality, and definitions of expert/research microscopy did not always match the quality criteria as used in this review. Highly similar quality criteria for microscopy have been used in a systematic review on the accuracy of malaria RDTs using microscopy and PCR as reference standards, with the difference that the reference standard was qualified as poor instead of unknown when the paper did not describe the use of a second independent microscopic or when the number of high power fields was not stated. In the present review, sufficient data were only available for real-time PCR to investigate whether existing heterogeneity could be explained by the quality of microscopy, which appeared not to be the case. To determine the usefulness of microscopy as a reference standard and to enable extensive subgroup analyses to be done, there is a need for detailed reporting on the quality of microscopy, describing the equipment and methods used, the number of microscopists, their independency, training level, and the procedure in case of discrepant results.

The low specificity of some molecular tools compared to reference standard microscopy may be explained by the fact that both types of tests aim to detect two different types of biomarkers. For example, when parasites adhere to endothelial cells (sequestration), microscopic detection may be complicated by low levels of parasites in the circulation, whereas this should be less problematic for (molecular) tests with a lower LoD. It can also be hypothesized that parasite DNA persists in blood during or after treatment and that therefore the detection of Plasmodium DNA does not necessarily indicate active infection, but might as well detect nuclear material from dead or drug-damaged parasites. However, Jarra and Snounou showed that total DNA clearance takes place between 24 and 48 h after injection of killed parasites into mice and that nuclear material from dead or drug-damaged parasites does not contribute significantly to amplification. This indicates that results from molecular tests reflect the presence of living parasites.

While molecular tools are reported to be very sensitive in detecting low parasite densities, there is some discussion about the clinical relevance of detecting such low parasitaemias, as these infections are often asymptomatic. However, they play an important role in transmission, especially in low prevalence areas. Furthermore, it has been found that the high sensitivity of PCR also leads to the identification of more patients with low parasitaemia in symptomatic microscopy-negative cases. The chances of feeling ill due to a sub-microscopic Plasmodium infection are therefore real, even though the presence of another cause of fever cannot be excluded. In this review, we did not assess the influence of parasite density on accuracy and additional (real-time PCR) data would be required to do so.

A possibly important confounder that was not taken into account in this review may be the input volume of blood used in the molecular assay, which is neither commonly described nor standardized. This is especially important for correctly diagnosing patients with low parasite densities, as chances of not capturing a parasite in a particular sample increase with decreasing parasitaemia. Recent research using high-volume real-time PCR (≥250 µl input material) shows a 50 times increase in analytical sensitivity compared to other common real-time PCR assays.

Sensitivity of a molecular test is also impacted by copy number, and stratification of data by DNA and total NA amplification is therefore of potential interest. For example, 18S rRNA contains about 3500 copies/parasite, which is much more than the commonly used 18S rRNA-coding genes (4–8 copies/parasite) or the mitochondrial cytochrome b (20–150 copies/parasite), implying that RNA potentially has a higher analytical sensitivity than DNA. Since only one study using reverse-transcriptase PCR was included, no such stratification could be done with the currently available data.

Prevalence was highly variable between studies (Table 1). This might have introduced bias in summary estimates of
sensitivity and specificity, as accuracy may vary with prevalence\textsuperscript{104}. The results of this review can therefore not be extrapolated to accuracy in screening asymptomatic populations or in febrile non-immune returning travelers.

Most studies were conducted in Asia (58\%) and Sub-Saharan Africa (28\%). No difference in accuracy between continents was found. There was high variability in the species detected by the index tests and no subgroup analysis could be done to determine whether a difference in accuracy exists, especially between \textit{P. falciparum} and \textit{P. vivax}. In line with the presence of both \textit{P. falciparum} and \textit{P. vivax} in many areas in the Asian-Pacific, most assays were multiplex systems and some presented data in their 2 × 2 tables for each species separately. However, in most studies too few samples per species were available to calculate sensitivity and specificity individually and therefore a joined estimate for all species together was given. Whether a difference in accuracy between species exists can only be established by studies designed to separate accuracy by species.

While some commercial real-time PCR kits are available, thus far they do not appear to be widely used. Most laboratories developed an in-house assay or used published protocols, often with small changes to the original method. Therefore, very little standardization of molecular malaria tests exists, complicating data comparison between laboratories\textsuperscript{18}. The results of this review indicate that especially for conventional and real-time PCR assays, many different protocols are in use. Studies evaluating nested PCR, on the other hand, refer in most cases to the method of Snounou et al.\textsuperscript{95,96} implying that some of the outliers of low sensitivity cannot be explained by differences in the original protocol used. Remarkably, especially for real-time PCR, the fact that many different protocols are in use did not seem to influence sensitivity. Still, standardization of commercial tests through FDA approval or the WHO prequalification program of \textit{in vitro} diagnostics would enhance the reliability and comparability of assays and would reduce the burden of quality assurance on individual laboratories\textsuperscript{105}. Additionally, even though accuracy of most molecular assays presented in this review is high, the experience of laboratory personnel is crucial in achieving this. Proficiency testing/External Quality Assessment (EQA) programs would provide insight in the quality of work of laboratories routinely performing molecular malaria diagnostics\textsuperscript{106,107}. For real-time PCR, guidelines are available presenting the minimum information for publication of quantitative real-time PCR experiments (MIQE), specifically addressing the potential to accurately quantify target nucleic acids\textsuperscript{108}. Their widespread use would increase experimental transparency and consistency between laboratories performing real-time PCR.

Implications for practice/further research

Even though it has been shown that PCR is highly sensitive and specific, the required infrastructure, trained personnel and costs hamper its implementation in resource-limited settings. LAMP and other novel molecular tools under development, as presented in this review, may be suitable alternatives to detect low-density infections. The different characteristics of these tests bring advantages for use in resource-limited settings, related to sample preparation, the amplification process and/or the read-out of the test.

Simplifying sample preparation reduces sample handling time and the number of required reagents. There are several assays that use whole blood or only require simple extraction methods. PCR-NALFIA can use whole blood as input material, whereby the need for sample preparation is completely avoided\textsuperscript{56}. The RNA-hybridization assay requires lysed red blood cells (RBCs)\textsuperscript{73} and for LAMP several crude extraction methods, like heat-treatment\textsuperscript{70,109} and boil-and-spin methods\textsuperscript{55} have been compared to commercial extraction kits. However, LAMP specificity appears to decrease when using simplified extraction methods (see Figure 7), indicating a need for improvement for use in field settings.

Using semi-isothermal amplification methods, LAMP, NASBA and the RNA-hybridization assay only need two temperatures for amplification, thereby circumventing the need for expensive PCR-machines\textsuperscript{110}. Instead, simple heat blocks or water baths can be used. However, the RNA-hybridization assay requires an overnight incubation step, which makes the assay in its current format unsuitable to be used in PoC settings, where results need to be produced as soon as possible in order to initiate proper treatment. When using LAMP or NASBA it would be possible to present results within hours, but NASBA has the disadvantage that RNA extraction is required, making it more suitable for research purposes than as a PoC test.

Traditional gel-electrophoresis read-out systems have the disadvantage that they need expensive equipment, produce hazardous waste and usually take at least 45 min. For LAMP, the read-out can be done visually using a UV-light, which makes the detection simple and extremely fast. The lateral flow device as a fast (10 min) read-out system in PCR-NALFIA is similar to the one used for RDTs, but detects amplicons instead of antibodies, enabling multiple species to be detected at the same time using a simple visual read-out. The detection system of PCR-ELISA is easy and fast, but requires an ELISA plate reader.

Costs are an important factor when evaluating whether a test is suitable for a particular purpose or not. This factor was not included and additional data would be required to get a complete overview of the costs per assay, including equipment, reagents, labor, training and maintenance. Currently available overviews mainly include equipment and reagent costs\textsuperscript{12,30,109} and do not cover all tests presented in this review.

The characteristics of a test that are most important in a certain situation are determined by its intended use (clinical practice or control/elimination programs) and setting (ranging from field conditions to reference laboratories)\textsuperscript{105,111,112}. For clinical practice, besides high accuracy, it is important that the test is fast, easy to perform on site and robust under adverse environmental conditions. Currently, LAMP and PCR-NALFIA are probably the most likely candidates to meet these characteristics. LAMP for its semi-isothermal amplification and PCR-NALFIA for using whole blood as input material. Both have an easy and fast read-out system. Preferably, no cold chain would be required to store reagents in final test formats. For the smaller clinics, high-throughput is less essential and tests will be done mostly for individual
(or small groups of) patients. Species differentiation may be especially important in areas where *P. vivax* is present and treatment regimens are based on the infecting species. Studies on asymptomatic populations were excluded from this review, but form an important subgroup for future reviews given the current focus on malaria screening and elimination programs. In such a review, it would be important to focus on the relation between clinical and analytical sensitivity, as the proportion of low-density infections is likely to be high in asymptomatic populations and parasite levels below the LoD of the test under evaluation will lead to false-negative test results. In these populations, poor analytical sensitivity is therefore a likely cause of poor clinical sensitivity. In the present review, studies evaluating clinical accuracy of the RNA-hybridization assay, LDR-FMA, PET-PCR and reverse transcriptase PCR specifically mention a possible application in screening programs in low transmission settings. For these programs, high-throughput and superior sensitivity are crucial. Preferably, the test should be able to discriminate between the five human infecting *Plasmodium* species. However, the delay caused by shipment and analyses of samples remains a challenge, as all positive cases need to be traced back and receive treatment as soon as possible. For elimination settings, portable equipment as described by Canier et al. may therefore be an alternative.

Due to the apparent implementation difficulties associated with molecular diagnostics, they are unlikely to replace microscopy or RDTs as first line diagnostics in remote settings in the near future. However, they may be a powerful alternative, especially in areas where prevalence is declining, which increases the need for more sensitive diagnostics, both in clinical practice and for screening asymptomatic populations. The further development of (near) PoC diagnostics for malaria could solve some of the implementation issues.

**Conclusions**

No important difference in accuracy was found between the most commonly used PCRs, with most of them having high sensitivity and specificity, implying that there are no reasons other than practical ones to choose one technique over the other. However, summary estimates should not be interpreted in isolation from the results of the original paper, as considerable heterogeneity exists, especially in the evaluations of conventional PCR. Alternative diagnostics appear to be highly accurate as well, with some offering potential to be used in resource-limited settings.

Future diagnostic accuracy studies should be reported according to Standards for the Reporting of Diagnostic Accuracy Studies (STARD) guidelines, to prevent inadequate reporting, maximize comparability and ensure inclusion in the meta-analysis. Direct comparisons of molecular tests (within-study) would facilitate stronger evidence of differences in accuracy between tests. Further research on the implementation of (near) PoC molecular diagnostics and their potential roadblocks is essential.

**Acknowledgements**

We would like to thank all authors who kindly provided additional data not published in the original papers.

**Declaration of interest**

This study was supported by EU FP7-Health-2013.0-1. Project no.: 601714. Translation of the direct-on-blood PCR-NALFIA system into an innovative near point-of-care diagnostic for malaria (DIAGMAL). J.R. is working on the development of the direct-on-blood PCR-NALFIA system. The authors alone are responsible for the content and writing of this article.

**References**


Supplementary material (Additional files 1 and 2) available online