Meta-analysis: accuracy of rapid tests for malaria in travelers returning from endemic areas


CRD summary
This review evaluated rapid diagnostic tests for ruling out malaria in non-immune travellers. The authors concluded that histidine-rich protein-2-based tests are a useful adjunct to microscopy in excluding Plasmodium falciparum infection in centres without major expertise in tropical medicine, but more research is required for other species. The review was generally well conducted and the conclusions should be reliable.

Authors' objectives
To determine the accuracy of rapid diagnostic tests for ruling out malaria in non-immune travellers returning from malaria-endemic areas.

Searching
MEDLINE, EMBASE, CAB Health and CINAHL were searched from 1988 to September 2004 without language restrictions; the search terms were reported. Conference proceedings and references of retrieved papers were also searched, while experts and manufacturers were contacted.

Study selection

Study designs of evaluations included in the review
Diagnostic accuracy studies were eligible for inclusion.

Specific interventions included in the review
Studies evaluating rapid diagnostic tests were eligible for inclusion. If more than one test was evaluated based on the same target antigen, preference was given to 3-band tests; if only 2-band tests were evaluated, preference was given to the test with the greatest number of patients, or the more recently developed test if the sample sizes were the same. Studies of self-administered diagnostic tests were excluded. The index tests evaluated in the included studies were 2-band histidine-rich protein-2 (HRP-2; ParaSight F, PATH IC Strip, ICT Pf, MAKROmed), 3-band HRP-2 (ICT Pf/Pv, NOW ICT Pf/Pv) and 3-band parasitic lactate dehydrogenase (LDL; OptiMal, OptiMal IT). Most of the studies were performed in regional or national reference centers for tropical diseases.

Reference standard test against which the new test was compared
Studies where the index tests were compared with microscopic examination or polymerase chain reaction (PCR) were eligible for inclusion. The combined results of PCR and microscopic examination were used as the reference standard of choice. The results of PCR were used in preference to microscopic examination where combined results were not provided. Where quantitative buffy coat was also used, patients with pure gametocytemia were considered positive. Six of the included studies used PCR as the reference standard and fifteen used microscopy, with six of these using additional reference tests in some patients.

Participants included in the review
Studies of non-immune individuals suspected of having malaria were eligible for inclusion. Studies where more than 10% of participants were immune were excluded. The studies were conducted in Europe, North America or Australia, with participants having travelled to Africa, Asia, or central and South America.

Outcomes assessed in the review
Studies had to present data which allowed 2x2 contingency tables to be constructed. A predefined negative likelihood ratio (LR) for Plasmodium (P.) falciparum was the primary measure of accuracy.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.
Assessment of study quality

Study quality was assessed in relation to study design, patient spectrum, performance of the reference standard for all participants, and blinding. Two reviewers independently assessed study quality. Any disagreements were resolved by consensus.

Data extraction

Two reviewers independently extracted the data. Any disagreements were resolved by consensus. Sensitivity, specificity and positive and negative LRs were calculated for each study.

Methods of synthesis

How were the studies combined?
The studies were grouped by targeted parasite and type of test, and pooled LRs were calculated using a random-effects meta-analysis. The results were plotted in receiver operating characteristic (ROC) space (sensitivity against 1-specificity). In the main analysis, infection with Plasmodium species other than P. falciparum were considered negative. In analyses of 3-band tests, negative LRs for infection with P. malariae, P. viva or P. ovale were calculated after excluding individuals infected with P. falciparum. Pooled negative and positive LRs were calculated for HRP-2 versus LDH using data from studies that directly compared these two tests for detecting P. falciparum.

For infection with P. falciparum, publication bias was investigated using funnel plots of negative LRs plotted against sample size.

How were differences between studies investigated?
Heterogeneity between the studies was assessed using the I-squared (I²) statistic, with I² less than 30% considered mild heterogeneity and I² greater than 50% considered pronounced heterogeneity. Subgroup analyses were conducted investigating the impact of antigen targeted by the index test, test generation (2-band or 3-band) and parasite density. A meta-regression was used to assess the impact of target antigen, quality criteria and sample size.

Results of the review

Twenty-one studies (n=5,747) were included in the review.

Only 2 studies met all five quality criteria. Six of the 21 studies were prospective, 10 enrolled consecutive patients, 9 reported blinding to the reference test results, 7 reported blinding to the index test results and all patients received a reference standard.

P. falciparum.

HRP-2 detection (19 studies): the sensitivities ranged from 88 to 99% and the specificities from 95 to 100%. The pooled negative LR was 0.08 (95% confidence interval, CI: 0.06, 0.10). There was no statistically significant difference in the negative LRs of 2- and 3-band tests, but the negative LRs were statistically significantly higher for second-generation 3-band tests than for first-generation 3-band tests (P=0.04). The 3-band tests had statistically significant, higher positive LRs than 2-band tests (34.7 versus 98.5, P=0.003). Performance was unsatisfactory for parasite densities up to 100 parasites/microL, good between 101 and 1,000 parasites/microL, and excellent for higher densities (5 studies).

Parasite LDH detection (6 studies): the sensitivities ranged from 79 to 95% and the specificities from 98 to 100%. The pooled negative LR was 0.13 (95% CI: 0.07, 0.22; heterogeneity was statistically significant (P<0.001). The 4 studies directly comparing HRP-2 and LDH detection found similar results (negative LR 0.06 versus 0.19, P=0.002).

There was no evidence of publication bias: the funnel plot of the negative LR was symmetrical and the P-value from the regression test was 0.76.

P. vivax.
HRP-2/aldose-based tests (5 studies): the sensitivities ranged from 46 to 93%, with high specificities. The negative LR was 0.24; heterogeneity was statistically significant. Second-generation 3-band tests had statistically significant, higher negative LRs than first-generation tests (P=0.001). Performance was unsatisfactory for parasite densities up to 1,000 parasites/microL (1 study).

Parasite LDH detection (6 studies): the sensitivities ranged from 62 to 95%, with high specificities. The negative LR was 0.13; heterogeneity was statistically significant.

P. ovale and P. malariae.

HRP-2 detection (4 studies): the sensitivities ranged from 7 to 80%, with high specificities. The negative LRs were close to 1; heterogeneity was statistically significant. Second-generation 3-band tests had statistically significant, higher negative LRs than first-generation tests (P=0.05).

Parasite LDH detection (5 studies): the sensitivities ranged from 36 to 95%, with high specificities. The negative LRs were close to 1; heterogeneity was statistically significant.

Authors’ conclusions
HRP-2-based rapid malaria tests were a useful adjunct to microscopy in excluding P. falciparum infection in centres without major expertise in tropical medicine, but more research is required for other species.

CRD commentary
The review question was clear in terms of the participants, index and reference tests, study design and outcome. Several relevant databases were searched, and attempts were made to minimise publication and language bias. The data extraction and quality assessment were carried out in duplicate, thus reducing the potential for error and bias during these stages of the review. However, the potential for selection bias remains.

Heterogeneity was adequately assessed using an appropriate statistic, and predefined subgroup analyses and meta-regression were used to examine the influence on the results of various factors. The plotting of the results in ROC space without a summary curve was appropriate given the heterogeneity between the studies. Most of the included studies were set in centres for tropical diseases and it was unclear if these results would generalise to other settings. The results were only consistent among studies for HRP-2-based tests for P. falciparum. Despite the potential for selection bias, the review was generally well conducted and the conclusions should be reliable.

Implications of the review for practice and research
Practice: The authors stated that antimalarial therapy and repeat testing may be appropriate after a negative HRP-2 test result if the suspicion of malaria persists, particularly if the patient’s condition deteriorates.

Research: The authors stated that further research should include randomised trials with clinical end points where patients with suspected malaria are randomly assigned to diagnostic strategies including and excluding the use of rapid testing. They also stated the need for more studies on the use of HRP-2/aldolase-based 3-band tests for the initial diagnosis of species other than P. falciparum, and on methods to improve the diagnostic accuracy of tests for P. ovale and P. malariae.

Funding
Swiss Federal Office of Public Health, grant number 03.000048.

Bibliographic details
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.