WHO selection criteria for procuring malaria RDTs

Working paper for discussion by the Malaria Policy Advisory Committee

WHO/GMP formulated the first recommendations on criteria for selection of malaria RDTs in 2010, based on the advice received by independent experts convened at a WHO Technical Consultation held in October, 2009. The data on comparative rapid diagnostic test (RDT) performance is based on the results of WHO Malaria RDT Product Testing Programme, a joint project of TDR, Foundation for Innovative New Diagnostics (FIND), US Centers for Disease Control and Prevention and WHO/GMP, involving collaboration with a number of research institutions and control programmes in malaria endemic and non-endemic countries. The WHO/GMP recommended selection criteria for procurement malaria RDTs form the basis for WHO RDT procurement practices, and are shared as an information note on the WHO/GMP website for use by WHO Member States and interested agencies.

WHO malaria rapid diagnostic test (RDT) performance evaluation

WHO currently runs an evaluation programme for malaria RDTs on which current WHO procurement recommendations and those of other agencies are based. This programme includes (1) the largest WHO-coordinated product testing programme for a health commodity, which recently completed its third round of testing, having evaluated and published detailed comparative data on 120 products since 2009, and (2) a lot-testing programme that has evaluated over 700 lots of malaria RDTs since 2008 and provides batch testing to country programmes on request prior to deployment and use in the field.

Figure 1: Response to WHO Malaria RDT Product Testing Expression of Interest: Rounds 1-4

---

3 WHO information note on recommended selection criteria for procurement of malaria rapid diagnostic tests (RDTs) [http://www.who.int/malaria/diagnosis_treatment/diagnosis/RDT_selection_criteria.pdf](http://www.who.int/malaria/diagnosis_treatment/diagnosis/RDT_selection_criteria.pdf)
Since 2010, lot-testing at a WHO-FIND quality assessed laboratory is mandatory for all procurement through the US President’s Malaria Initiative (PMI), and is required by the Global Fund. In the first 3 quarters of 2011 300 lots were evaluated and capacity can easily be expanded. Currently WHO and FIND do not charge fees to manufacturers, programmes or Agencies submitting their product for evaluation through either of these programmes.

**Figure 2:** Lot testing trends 2007-2010

The same programme currently provides reference parasite panels for research and development to manufacturers, and is developing recombinant antigen-based panels to form the next generation of reference materials for malaria RDT evaluations that will allow for product testing at much lower costs as well as standardized country-based lot-testing.

Immuno-chromatographic tests are multi-component biological tests in which the performance may be significantly affected by a large number of variables, which can result in lot-to-lot variation. The current WHO product testing programme is producing detailed comparative performance data on a high number of products (120 have been assessed, including 23 products assessed in Round 3 that had been previously assessed in Round 1).

**Figure 3:** Improvement in RDT performance: Results of re-submitted products: Round 1 (2009) and Round 3 (2011)

**Figure 1:** Improvement in RDT Performance: Results of resubmitted RDTs: Round 1(2009) and Round 3 (2011)
The lot-testing programme is conducting batch testing prior to release to the field, to ensure performance at release irrespective of manufacturing conditions. Both programs have a clear impact on the quality of RDTs being procured for public sector use: recent FIND market survey data indicates that in 2010, 78% (~78 million) of RDTs manufactured met the most stringent WHO procurement criteria, compared with just 23% (~6 million) in 2007. In parallel to these trends, the frequency of lot testing failures has progressively declined over the years. In 2010, batches had a 100% pass rate; only one failure was seen during the first half of 2011, indicating that manufacturers are maintaining quality, at least when they are aware the lots will be evaluated.

**WHO recommended selection criteria for procurement**

In October 2009 WHO convened a technical consultation to review the evidence base for thresholds of diagnostic performance required by current malaria diagnostic tests and to make recommendations on their use. The consultation reviewed the clinical significance of parasite densities in patients with uncomplicated *P. falciparum* and *P. vivax* malaria, the risks of missing low parasite densities with routine field microscopy and most RDTs on the market, the implications of results of the product testing programme for malaria RDTs to provide advice to procurement agencies.

The Consultation reviewed factors affecting the frequency of low-density infections, including host immunity, parasite factors, stage of illness and effectiveness of treatment, and focussed on the frequency of parasite densities < 200/µL in patients seeking treatment in health facilities and its relation with transmission intensity and parasite species.

In high-transmission areas, only about 5% of patients with *P. falciparum* malaria have parasite densities < 200/µL. In low-to-moderate transmission areas, 5–10% of patients with *P. falciparum* malaria have parasite densities < 200/µL. Patients with *P. vivax* malaria present with parasite densities < 200 per microliter more commonly than those with *P. falciparum* malaria (~15%). The frequency of low parasite densities (<200/µL) is higher in population and household surveys than among symptomatic patients who present to health facilities for treatment.

Based on these considerations and the review of results product testing and lot testing of tests on the market, the participants in the 2009 consultation recommended the following selection criteria for RDT procurement:

A. The *P. falciparum* panel detection score for high transmission areas should be at least 50% at 200 parasites/µL. Since the extent of high transmission areas is likely to decrease with effective malaria control, a panel detection score well above this level should become the basis for product selection in the future years.

---

4 WHO Malaria RDT Product Testing: Panel detection score ≥75% for panels of *P. falciparum* and/or *P. vivax* at 200 parasites/µL

5 See ANNEX 1 for a brief discussion on current status of WHO system for prequalification of malaria RDTs

6 The term ‘panel detection score’ (PDS) is a composite index of test positivity as well as of inter-test and inter-lot consistency and is not a measure of clinical sensitivity.

7 ‘High transmission’ areas are hyperendemic and holo-endemic areas in which the prevalence rate of malaria is over 50% during most time of the year among children from 2 to 9 years old. In these areas by late infancy or early childhood practically all individuals are infected.
B. The *P. falciparum* panel detection score for low⁸ and moderate⁹ transmission areas should be well above 50% at 200 parasites/μL (e.g. ≥75%).

C. The *P. vivax* panel detection score for low and moderate transmission areas should be equivalent to those for *P. falciparum* - well above 50% at 200 parasites/μL (e.g. ≥75%).

D. In all areas false positive rate should be less than 10%.

E. In all areas invalid rate should be less than 5%.

Based on the above criteria out of the 95 unique¹⁰ RDTs assessed in Rounds 1-3, a total of 24 Pf-only RDTs meet the above criteria for use in high transmission areas, and 21 Pf-only RDTs, 13 combination RDTs, 2 pan RDTs and 1 Pv-only RDT meet the criteria for areas with low or moderate transmission.

**Calculation of the Panel Detection Score**

The panel detection score ('detection rate' in the WHO/FIND round 1 evaluation) is a number between 0 and 100, calculated as the proportion of times a malaria RDT gives a ‘pass’ result in all tests on both lots tested in multiple samples of parasite panels of wild type parasites at a specific parasite density, i.e. four tests at 200 parasites/μL and two at 2000 parasites/μL. In each round, the panel detection score at low parasite densities was calculated against panels derived from 79-100 samples of *P. falciparum* and 20-40 samples of *P. vivax*. Invalid tests are excluded from the analysis. In the calculation of the score for low parasite densities, all four tests (two each from two different production lots) should be positive in order for the test to ‘pass’. In the example shown in the figure, the test ‘fails’ to detect parasite in a given sample if three of four tests are positive.

**Figure 4**: Determination of WHO Product Testing panel detection score at low parasite density (200 parasites/μL)

---

⁸ 'Low transmission' areas are hypo-endemic areas in which the prevalence rate of malaria is 10% or less during most time of the year among children from 2 to 9 years old. Here a person may attain adolescence before malaria infection is acquired and may escape acquiring a malaria infection altogether.

⁹ 'Moderate transmission' areas are meso-endemic areas in which the prevalence rate of malaria is 11-50% during most time of the year among children from 2 to 9 years old. Here the maximum prevalence of malaria infection occurs in childhood and adolescence, though still not unusual for adult life to be attained before acquiring infection.

¹⁰ This excludes 2 products that did not pass Phase 1 and 23 products resubmitted under the same product codes.
The panel detection score is different from the sensitivity or positivity rate, as it includes a measurement of intra-lot consistency and inter-lot variation. Thus, a PDS of 80% at a parasite density of 200/µL is a good result and does not correspond to a sensitivity of 80% observed in the field. The largest difference in test performance that allows differentiation of RDTs that perform well and those that perform poorly is reflected in the panel detection score at the lower parasite density (200 parasites/µL).

**Relation between Panel Detection Score and Sensitivity**

The diagnostic performance of malaria RDTs, as measured from the panel detection score may not be directly related to the sensitivity of the test in clinical testing. More specifically, in product testing parasitized blood samples from patients are diluted to ensure they consistently have the same parasite density (and range of antigen concentrations); however in the field, samples of parasitized blood from patients are much more likely to have heterogeneous parasite densities -- generally with parasitaemias higher than 200 parasites/µL.

The performance of malaria RDTs can also be assessed from their diagnostic sensitivity and specificity in target populations, as reported in the scientific literature. However, the quality of studies is variable, and the reported parameters depend closely on samples selected for the study, RDT quality and storage conditions, the user’s skill in preparing and interpreting test results and the quality of the microscopy used as reference standard. A Cochrane review of *P. falciparum* RDT field performance has recently been published.

The series of factors which may affect performance testing in a laboratory setting compared to field trials and may explain discrepancies in performance (panel detection score) in WHO RDT Product Testing and (populations based) sensitivity are listed in the table below.

---

Table 1: Reasons for discrepancy between panel detection score and clinical sensitivity

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure to extreme temperatures</td>
<td>High temperatures accelerate degradation (deconjugation of the signal antibody–dye complex, detachment of capture antibody from the wick, and change the binding sites of antibodies and the nitrocellulose). Freeze-thawing may have similar effects.</td>
</tr>
<tr>
<td>Age and storage of blood sample</td>
<td>Stored blood may lose antigen activity; early lysis and protein coagulation can inhibit flow. The rate of loss of antigen activity varies among antigens. Lysis of cells can occur during mixing and storage.</td>
</tr>
<tr>
<td>Preparation of dilutions</td>
<td>Cell lysis and aggregation of parasitized cells can affect flow.</td>
</tr>
<tr>
<td>Visual acuity of technician</td>
<td>Can affect reading of faint test lines at low parasite density.</td>
</tr>
<tr>
<td>Patient and parasite</td>
<td>Parasite density affects sensitivity.</td>
</tr>
<tr>
<td></td>
<td>Parasite density and parasite load (including sequestered parasites) determine antigen levels.</td>
</tr>
<tr>
<td></td>
<td>Antigen production varies during the parasite life cycle and between parasite strains. Previous treatment and its effectiveness varies among patients. Factors that cause false-positive results can vary among patients. Antigen activity may be different in wild and cultured parasites.</td>
</tr>
<tr>
<td>Reference standard (microscopy or PCR)</td>
<td>Poor sensitivity reduces apparent RDT specificity.</td>
</tr>
</tbody>
</table>

Small differences in panel detection scores among the better-performing RDTs in an evaluation are unlikely to result in noticeable differences in clinical sensitivity. On the other hand, the panel detection score at 200 parasites/µL provides an indication of which products are likely to be more sensitive in the field, particularly in populations with low-density infections.

Re-visiting WHO procurement criteria for malaria RDTs

The WHO recommendations set in 2009 on selection criteria for procurement of malaria RDTs have been considered by some stakeholders (e.g. procurement and funding agencies) as setting the bar too low, particularly the recommended threshold of *P. falciparum* panel detection score (PDS) for high transmission areas at 50% at 200 parasites/µL. However, other stakeholders (e.g. manufacturers and some end-users) have concerns that the current bars, particularly for combination tests, are too high and exclude tests that perform well in field settings.

It is almost certain that some of the concerns that the bar is too low stem from equating PDS with sensitivity, thereby implying that WHO condones detecting (and therefore treating) just 50% of patients with potentially fatal *P. falciparum* malaria. This however, is a very flawed conclusion. As previously mentioned, in reality, it is estimated that only 5% of the population in a high transmission zone would have parasite densities <200/µL and of these 50% could be missed, based on the current procurement criteria. It is on these grounds, plus the limited number of *P. vivax* samples included in the WHO Product Testing protocol and the Programme’s rigorous requirements for inter-test and inter lot consistency, that some manufacturers think the current criteria may be too stringent.
With arguments on both sides, it is clear that a change in the current criteria must be accompanied by reasonable evidence of harm associated with the current criteria or conversely evidence/predicted public health benefits of raising the performance requirements.

**Considerations in favour of an increase to a PDS of 75% as a threshold:**

An increase in the PDS to 75% for Pf-tests in high-endemic areas will align the threshold used for this setting with that already used for the other settings (Pf-tests in low to moderate-endemic areas, and Pv-tests for any level of endemicity). This alignment will render RDT selection by countries much easier as they will not need to take into account local transmission which is changing in time and space.

An increase in the PDS to this new threshold of 75% will be met by 21 Pf-only RDTs, as opposed to 24 Pf-only RDTs if the threshold remains unchanged for areas of high transmission.

Many of the 23 tests which were re-submitted from Round 1 to Round 3 have been found to have increased panel detection scores, with the highest improvement in mean/median values for *P. vivax* scores (see figure below).

The shift of the threshold of panel detection score from 50% to 75% will be in line with conclusions of experts convened by WHO Technical Consultation held in 2009, which recommended that “As malaria control improves, there will be greater demand for RDTs that consistently have detection rates of at least 75% at low densities (200 parasites per microlitre) of *P. falciparum* and *P. vivax* parasites.”

**Figure 5:** Improvement in Panel Detection Score in re-submitted RDTs products between Round 1 (2009) and Round 3 (2011)
Considerations against an increase to a PDS of 75% as a threshold:

The relationship between panel detection score and clinical sensitivity will vary depending on the local epidemiology, and small differences in PDS may not have relevance in terms of clinical impact. Indeed, several studies have shown that the use of RDT for clinical management is safe, even in moderate endemic areas when using RDTs with *P. falciparum* panel detection score (PDS) at 200 parasites/µL of much less than 75% (see Table 2).

The distribution of PDS results against panels of wild type panels of both *P. falciparum* and *P. vivax* diluted at low parasite densities (200 parasites/µL) is linear with small incremental differences making any threshold level arbitrary, possibly unfair and probably clinically irrelevant (no public health impact).

If a new threshold of 75% were adopted, 3 Pf-only RDTs will be no longer eligible for procurement for areas of intense transmission; the impact of this in terms of market share and use is unknown. However, since the Round 4 of WHO Product Testing is on-going, it is not yet know how many RDTs and manufacturer could be potentially affected by an increased in the recommended threshold for procurement of malaria RDTs.
Figure 6: *P. falciparum* PDS at 200 parasites/μL for RDTs tested during Rounds 1, 2 and 3. Red stars show 3 RDTs which will be not eligible for procurement for areas of high transmission if the recommended threshold were increased from 50% PDS (blue solid line) to 75% (blue dotted line).
<table>
<thead>
<tr>
<th>Reference</th>
<th>Country, year of patient inclusion</th>
<th>- Age group</th>
<th>Study design</th>
<th>RDT product used and Pf PDS at 200 parasites/µL (Round1)</th>
<th>Adherence to negative RDT result (no antimalarial given) in intervention arm</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Msellem et al, <em>PLoS Medicine</em> 2009</td>
<td>Zanzibar, 2005</td>
<td>All ages (55% were children under five)</td>
<td>Cross-over randomized control trial of dispensaries using RDT (level of randomization: dispensary)</td>
<td>ParaCheck® Pf</td>
<td>100%</td>
<td>No initially negative case developed severe malaria&lt;br&gt;Deaths: None&lt;br&gt;No reattendance within 14 days: 97% in intervention arm and 95% in control arm</td>
</tr>
</tbody>
</table>
| D'Acremont et al *Clinical Infectious Diseases* 2010 | Tanzania, 2007-2008 | Children under five | Dispensaries using RDT No control arm | ParaHit® Pf | 100% (per protocol) | No initially negative case developed severe malaria<br>Deaths: 2 (1 severe sepsis and 1 severe pneumonia)
* Hospitalizations: 0.5%
* Clinical clearance at day 7: 97% |

- Desired outcome: number of initially negative patients who developed severe malaria
- Other valid outcomes:
  - Complications: hospitalizations; deaths
  - Clinical outcome at follow-up: clearance of fever (by history and/or elevated temperature); clearance of all symptoms; absence of reattendances

---

12 Both children were negative at inclusion. One developed severe sepsis and one severe pneumonia; they were both tested negative again by RDT and microscopy.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Country, year of patient inclusion</th>
<th>Age group</th>
<th>Study design</th>
<th>RDT product used</th>
<th>Adherence to negative RDT result (no antimalarial given) in intervention arm</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeboah et al, Plos Medicine 2010</td>
<td>Zambia, 2008</td>
<td>Children under five</td>
<td>Randomized control trial of CHW using RDT and respiratory rates (level of randomization: CHW)</td>
<td>ICT diagnostics® Pf, Pf PDS = 82.3%</td>
<td>99.6%</td>
<td>Desired outcome: number of initially negative patients who developed severe malaria</td>
</tr>
<tr>
<td></td>
<td>PR 28% P falciparum</td>
<td>Intervention arm: 1017 Control arm: 2108</td>
<td></td>
<td></td>
<td></td>
<td>Other valid outcomes: hospitalizations; deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clinical outcome at follow-up: clearance of fever (by history and/or elevated temperature); clearance of all symptoms; absence of reattendances</td>
</tr>
<tr>
<td>Tiono et al, In preparation 2011</td>
<td>Burkina faso, 2009</td>
<td>Children under five</td>
<td>Randomized control trial of CHW using RDT and respiratory rates (level of randomization: village)</td>
<td>First Sign®, PDS = 31.6%</td>
<td>96.3%</td>
<td>No initially negative case developed severe malaria</td>
</tr>
<tr>
<td></td>
<td>PR 74% P falciparum</td>
<td>Intervention arm: 525 Control arm: 576</td>
<td></td>
<td></td>
<td></td>
<td>Deaths: none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Elevated temperature clearance: 100% in intervention arm and 99 in control arm</td>
</tr>
</tbody>
</table>

Both children were negative at inclusion. One developed severe pneumonia and was tested negative again by RDT; the other one developed severe gastro-enteritis but was not retested for malaria.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Country, year of patient inclusion</th>
<th>Age group</th>
<th>Study design</th>
<th>RDT product used</th>
<th>Adherence to negative RDT result (no antimalarial given) in intervention arm</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anyorigiya et al, In preparation 2011</td>
<td>Ghana, 2009</td>
<td>Children under five</td>
<td>Randomized control trial of CHW using RDT and respiratory rates (level of randomization: village)</td>
<td>ParaCheck® Pf PDS = 54.4%</td>
<td>96.7% No initially negative case developed severe malaria Deaths: none Elevated temperature clearance: 99% in intervention arm and 98% in control arm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PR 84% P. falciparum</td>
<td>Control arm: 591</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children under five Intervention arm: 584</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control arm: 591</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senn et al, In preparation 2011</td>
<td>Papua New Guinea, 2006-2010</td>
<td>Children less than 2 years</td>
<td>Dispensaries using RDT No control arm</td>
<td>ICT diagnostics® Combo PDS = 86.1%</td>
<td>100% (per protocol) Deaths: 3 (1 severe malaria + pneumonia, 1 severe pneumonia, 1 meningitis) Hospitalization and/or severe illness: 0.5% No reattendance within 7 days: 96%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PR 30% P. falciparum (19%), mixte (37%), non-falciparum (44%)</td>
<td>Children less than 2 years</td>
<td>Dispensaries using RDT No control arm</td>
<td>ICT diagnostics® Combo PDS = 86.1%</td>
<td>100% (per protocol) Deaths: 3 (1 severe malaria + pneumonia, 1 severe pneumonia, 1 meningitis) Hospitalization and/or severe illness: 0.5% No reattendance within 7 days: 96%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GM Pf: 22,196 p/µl GM Pv: 4,792 p/µl (range 40 - 654960)</td>
<td>Children less than 2 years</td>
<td>Dispensaries using RDT No control arm</td>
<td>ICT diagnostics® Combo PDS = 86.1%</td>
<td>100% (per protocol) Deaths: 3 (1 severe malaria + pneumonia, 1 severe pneumonia, 1 meningitis) Hospitalization and/or severe illness: 0.5% No reattendance within 7 days: 96%</td>
<td></td>
</tr>
</tbody>
</table>

14 One child was positive at inclusion and developed severe pneumonia. Two children were negative at inclusion; one developed severe pneumonia and one got meningitis; they were both tested negative again by RDT and microscopy.
WHO Prequalification of malaria RDTs

WHO Essential Health Technologies (EHT) Department has started approximately 5 years ago a pre-qualification (PQ) programme diagnostic devices for a number of diseases including malaria. This programme has based its evaluation scheme on the model for prequalification of medicines, including assessment of a product dossier, and inspection of the manufacturing site of each product. Manufacturers have received communication informing them that they should participate in the PQ programme, and several have submitted dossiers to date, including 37 malaria RDTs (as of 28 November 2011). So far 2 RDTs have been prequalified by this programme (of which one in dipstick format, no longer in use by malaria programs). WHO EHT applies a non-refundable assessment fee of US $12,000 to manufacturers submitting their product dossier for evaluation by WHO PQP. Since September 2010 the WHO/EHT programme has included the results of the WHO Product Testing Programme as a third evaluation component required to achieve full prequalification, in addition to dossier review and inspection of the manufacturing facilities. In addition, recently applications have been closed due to product testing results that do not meet current WHO procurement criteria requirements.

There are currently over 60 manufacturers of malaria RDTs and approximately 200 malaria RDT products commercially available, with a high rate of entry of new and modified products. The proposed PQ model for malaria RDTs demands significant time investments as in the case for medicines, due to the requirements of product dossier compilation, acceptance/review, correspondence on observations, inspection and reporting, addressing observations, and review of corrective actions/re-inspection for possible approval. Often manufacturers of RDTs rely on multiple manufacturing production facilities, and each of these would require separate inspections and certification by the PQ team, to provide prequalification status. Furthermore, manufacturers often make minor changes to their products to improve performance and operational characteristics. WHO/EHT requires that they be informed and that a detailed description/report of any product variations be provided. A decision is taken, on a case-by-case basis as to whether or not the change constitutes a new product and would require re-submission to PQP. This approach proves challenging in being able to respond in a timely manner to the current rate of new product entry (and variations) on the market; by the time one product is processed, it is likely that new and improved products will be entering the market.

A meeting of experts, external stakeholders (UN agencies, global health initiatives, national regulatory authorities, and NGOs) and Regional Offices was held on 4-6 Oct 2011 to review WHO prequalification. They proposed a reorganization of prequalification in WHO to consolidate several of the different programmes, and strengthen the links between prequalification and capacity building. The immediate need is for a technical review of the mechanism for prequalification of diagnostics in WHO.